CHAPTER 4

Optical Sensors and Actuators

The eye

The human eye, like that of other vertebrates, is a marvelous, complex sensor allowing us to perceive the world around us in minute detail and true colors. In fact, the eye is akin to a video camera. It consists of a system of lenses (the cornea and crystalline lens), an aperture (iris and pupil), an image plane (retina), and a lens cover (eye lids). In humans and animals that prey, the eyes point forward to create binocular vision with excellent depth perception. Many prey animals have side-facing eyes to increase their field of view, but the vision is monocular and lacks perception of depth. The eyelids, in addition to protecting the eye, also keep it clean and moist by distributing tears as well as lubricants (the conjunctiva) and protect it from dust and foreign objects in conjunction with the eyelashes. The front dome of the eye is made of the cornea, a clear, fixed lens. This is a unique organ, as it has no blood vessels and is nourished by tears and the fluid inside the eye sphere. Behind it is the iris, which controls the amount of light that enters the eye. On the periphery of the iris there is a series of slits that allow fluid to pass out from the eye sphere. This passes nutrients to the front of the eye and relieves the pressure in the eye (when this is not perfectly regulated one has glaucoma, a condition that can affect the retina and eventually can cause blindness). Behind is the crystalline lens, an adjustable lens that allows the eye to focus on objects as close as about 10 cm and as far as infinity. The lens is controlled by the ciliary muscle. When this muscle loses some function, the ability of the lens to focus is impaired, leading to the need for corrective action (glasses or surgery). The lens itself can cloud over time (cataracts), a condition that requires replacement of the lens. At the back of the eye lies the optical sensor proper—the retina. It is made of two types of cells: cone cells that perceive color and rod or cylindrical cells that are responsible for low-light (night) vision. The cone cells are divided into three types, sensitive to red, green, and blue light, with a total of about 6 million cells, most of them in the center of the retina (the macula). Rod cells are distributed mostly on the peripheral parts of the retina and are responsible for low-light vision. They do not perceive color but are as much as 500 times more sensitive than cone cells. There are also many more rod cells—as many as 120 million of them. The retina is connected to the visual cortex in the brain through the optical nerve. Although the lens of the eye is adjustable, the size of the optical ball also plays a role in vision. Individuals with larger eyeballs are nearsighted, those with smaller eyeballs are farsighted.

The sensitivity of the human eye ranges from roughly 10^{-6} cd/m^2 (dark night, rod-dominated vision, essentially monochromatic) to about 10^6 cd/m^2 (bright sunlight, cone-dominated vision, full color). This is a vast dynamic range (120 dB). The spectral sensitivity of the eye is divided into four partially overlapping zones. Blue cones are sensitive between about 370 nm and 530 nm, with peak sensitivity at 437 nm; green cones between 450 and 640 nm, with peak sensitivity at 533 nm; and red cones between 480 and 700 nm, with peak sensitivity at 564 nm. Rods are sensitive between about 400 and 650 nm, with peak at 498 nm. This peak is in the blue-green range. For this reason, low-light vision tends to be dark green.

It should also be noted that the human-type eye, a structure shared by many animals, is not the only type of eye. There are some 10 different structures ranging from simple light-sensitive cells that allow the organism to detect light but not to create images, to compound eyes made of thousands of simple, individual "eyes" particularly suited to detect motion but can only create "pixilated" images.

4.1 | INTRODUCTION

Optics is the science of light and light is an electromagnetic radiation that manifests itself either as an electromagnetic wave or as photons (particles with quanta of energy). Before continuing it is worth mentioning that the term light refers specifically to the visible spectrum of electromagnetic radiation as perceived by the human eye (see **Figure 4.1**), but because both below and above this spectrum the behavior of radiation is similar, the term light is normally extended to include a much wider spectrum that includes infrared (IR) radiation (below the frequency of visible light or "below red") and ultraviolet (UV) radiation (above the visible range or "above violet"). Even the nomenclature has been modified and we sometimes say IR light or UV light. These terms are incorrect but are in widespread use. The range that is properly called light is defined by the response of the human eye between 430 THz and 750 THz (1 THz) = 10^{12} Hz). In characterizing light it is more common to use wavelength, defined as the distance in meters the light wave propagates in one cycle or $\lambda = c/f$, where c is the speed of light and f is its frequency. The range of wavelengths in the visible light region is between 700 nm (deep red) and 400 nm (violet). However, the ranges of IR and UV radiation are not as well defined and, as can be seen in **Figure 4.1**, the lower range of IR radiation overlaps the higher range of microwave radiation (sometimes this upper range is called millimeter wave radiation), whereas the upper reaches of UV radiation reach into the X-ray spectrum. For the purpose of this discussion, the IR range is between 1 mm and 700 nm and the UV range is between 400 nm and 1 nm. What unifies this wide range for the purpose of this chapter is the fact that the principles of sensing are similar and based on essentially the same effects. It should also be pointed out that the term radiation here means electromagnetic radiation.

Optical sensors are those sensors that detect electromagnetic radiation in what is generally understood as the broad optical range—from far IR to UV. The sensing methods may rely on direct methods of transduction from light to electrical quantities such as in photovoltaic or photoconducting sensors or indirect methods such as

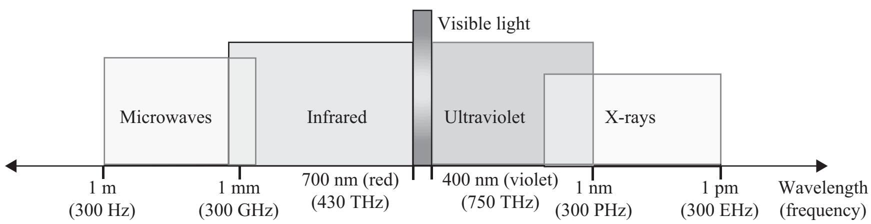


FIGURE 4.1 ■ Spectrum of infrared, visible, and ultraviolet radiation.

conversion first into temperature variation and then into electrical quantities such as in passive IR (PIR) sensors and bolometers.

There is a third method of sensing related to optics—sensors based on light propagation and its effects (reflection, transmission refraction), which will not be discussed here because the optical aspect is usually not the sensing mechanism, but rather an intermediate transduction mechanism. Nevertheless, the physics will be mentioned briefly for completeness.

4.2 | OPTICAL UNITS

The units used in optics seem to be more obscure than most. Thus it is useful to discuss these at this point. First, the SI units provide for a measure of **luminous intensity**, the candela (cd) (see **Section 1.6.1**). The candela is defined as the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} Hz and that has a radiation intensity of 1/683 W/sr. In short, the candela is a measure of radiation intensity.

Other units are often used. The **lumen** (luminous flux) is a candela steradian (cd·sr) and is a measure of power. The **lux** (a measure of illumination) is a candela steradian per square meter (cd·sr/m^2) and is therefore a power density. These are summarized in **Table 4.1**.

TABLE 4.1 ■ Optical quantities and their units

Quantity	Name	Unit	Derived unit	Comments
Luminous intensity	Candela	[cd]	[W/sr]	Power radiated per steradian
Luminous flux	Lumen	[cd·sr]	[W]	Power radiated
Illuminance	Lux	[cd·sr/m^2]	[W/m^2]	Power density
Luminance	Candela per meter square	[cd/m]^2	[W/sr·m^2]	Density of luminous intensity

EXAMPLE 4.1 Conversion of optical units

A point source emits uniformly in all directions in space (e.g., the sun may be considered a point source when viewed from Earth). Given a total power radiated of 100 W, calculate the source's luminous intensity and illuminance at a distance of 10 m from the source.

Solution: Since there are 4π solid unit angles in a sphere, the luminous intensity is

$$\text{luminous intensity} = \frac{100}{4\pi} = 7.958 \text{ W/sr}.$$

Since a candela is 1/683 W/sr, the luminous intensity is

$$\text{luminous intensity} = \left(\frac{100}{4\pi}\right) \times 683 = 7.958 \times 683 = 5435.14 \text{ cd}.$$

Although the units of illuminance are cd·sr/m^2, it is best to start from the power radiated. That power is spread over the sphere of radius $R = 10$ m, so we get

$$\text{Illuminance} = \frac{100}{4\pi R^2} = \frac{7.958}{10^2} = 0.0796 \text{ lux}.$$

It should be noted that whereas the luminous intensity is a fixed value that only depends on the source, illuminance depends as well on the distance from the source.

Note: A uniformly radiating source is called an isotropic source.

4.3 MATERIALS

The sensors/actuators discussed in **Chapter 3** and those that will follow take advantage of many physical principles. But, in addition, they take advantage of specific material properties, either of elements, alloys, or in other forms available, including synthetic and naturally occurring salts, oxides, and others. As we will discuss some of these, especially in conjunction with semiconducting materials, it is perhaps useful to bear in mind the periodic table (see the inside back cover). Many of the properties of materials are not specific to a single element, but rather belong to a group (often a column in the table of elements), and one can expect that if an element in a specific column is used for a given purpose, other elements from the same column may have similar properties and be equally useful. For example, if potassium (alkali column I) is useful in the production of cathodes for photoelectric cells, then lithium, sodium, rubidium, and cesium should also be useful. But there are clear limits. Hydrogen and francium, which are also in the same column, are not useful. The first because it is a gas, the second because it is radioactive. Similarly, if gallium-arsenide (GaAs) makes a useful semiconductor, so should indium-antimonide (InSb), and so on. We already saw some of these principles in discussing thermocouples. The elements in the VIII column—nickel, palladium, and platinum are used for various types of thermocouples together with elements from the IB and IIB columns. We shall refer to the periodic table often, but will also refer to many simple or complex compounds with specific properties that have been found to be useful in sensors and actuators. Here we will be concerned primarily with semiconductors, but other materials will become important in subsequent chapters.

4.4 EFFECTS OF OPTICAL RADIATION

4.4.1 Thermal Effects

The interaction of light (radiation) with matter results in absorption of energy in two distinct ways. One is thermal and is usually viewed as absorption of electromagnetic waves. The other is a quantum effect. The thermal effect is based on electromagnetic energy absorbed by the medium and converted into heat through the increased motion of atoms. This heat is sensed and translated into a measure of the incident radiation. Here we will not go beyond the understanding that by raising the temperature of a material, its electrons gain kinetic energy and may be released given sufficient energy and, of course, that this interaction can be used for sensing.

4.4.2 Quantum Effects

4.4.2.1 The Photoelectric Effect

The second effect is a quantum effect and is governed by photons, the particle-like manifestation of radiation. In this representation of light, and in general radiation, energy travels in bundles (photons) whose energy is given by Plank's equation:

$$e = hf, \tag{4.1}$$

where $h = 6.6262 \times 10^{-34}$ J/s or 4.1357×10^{-15} eV, which is Planck's constant, and f is the frequency in hertz. e is the photon energy and is clearly frequency dependent. The higher the frequency (the shorter the wavelength), the higher the photon energy. In the quantum mode, energy is imparted to materials by elastic collision of photons and electrons. The electrons acquire energy and this energy allows the electron to release itself from the surface of the material by overcoming the **work function** of the substance. Any excess energy imparts the electron kinetic energy. This theory was first postulated by Albert Einstein in his photon theory, which he used to explain the photoelectric effect in 1905 (and for which he received the Nobel Prize). This is expressed as

$$hf = e_0 + k, \tag{4.2}$$

where e_0 is the work function and is the energy required to leave the surface of the material (see **Table 4.2**). The work function is a given constant for each material. $k = mv^2/2$ represents the maximum kinetic energy the electron may have outside the material. That is, the maximum velocity electrons can have outside the material is $v = \sqrt{2k/m}$, where m is the mass of the electron.

A photon with energy higher than the work function will, in principle, release an electron and impart a kinetic energy according to **Equation (4.2)**. But does in fact each

TABLE 4.2 ■ Work functions for selected materials

Material	Work function [eV]
Aluminum	3.38
Bismuth	4.17
Cadmium	4.0
Cobalt	4.21
Copper	4.46
Germanium	4.5
Gold	4.46
Iron	4.4
Nickel	4.96
Platinum	5.56
Potassium	1.6
Silicon	4.2
Silver	4.44
Tungsten	4.38
Zinc	3.78

Note: 1 eV = 1.602×10^{-19} J.

photon release an electron? That depends on the quantum efficiency of the process. Quantum efficiency is the ratio of the number of electrons released (N_e) to number of photons absorbed (N_{ph}):

$$\eta = \frac{N_e}{N_{ph}}. \tag{4.3}$$

Typical values are around 10%–20%. This simply means that not all photons release electrons.

Clearly, for electrons to be released, the photon energy must be higher than the work function of the material. Since this energy depends on frequency alone, the frequency at which the photon energy equals the work function is called a **cutoff frequency**. Below it quantum effects do not exist (except for tunneling effects) and only thermal effects are observed. Above it, thermal and quantum effects are present. For this reason, low-frequency radiation (IR in particular) can only give rise to thermal effects, whereas at high frequencies (UV radiation and above) the quantum effect dominates.

This then describes the photoelectric effect, which is the basis for a number of sensing methods, as we shall discuss next. In all of these methods, surface electrons are released.

EXAMPLE 4.2 Minimum wavelength for photoelectric emission

Consider a photoelectric device intended for light detection.

a. Assuming it is made of a potassium-coated surface, what is the lowest wavelength that the device can detect?
b. What is the kinetic energy of an emitted electron under red light radiation at a wavelength of 620 nm?

Solution:

a. The photon energy is given in **Equation (4.2)**. With the photon energy equal to the work function we have

$$hf = e_0 \rightarrow f = \frac{e_0}{h} \quad \text{[Hz]}.$$

Since photons travel at the speed of light, the frequency may be written as

$$f = \frac{c}{\lambda} \quad \text{[Hz]},$$

where c is the speed of light and λ is the wavelength. The longest wavelength detectable is

$$\lambda = \frac{ch}{e_0} = \frac{3 \times 10^8 \times 4.1357 \times 10^{-15}}{1.6} = 7.7544 \times 10^{-7} \text{ m}.$$

This is 775.44 nm. From **Figure 4.1**, this is in the very near IR region.

b. At 620 nm, the frequency is c/λ and from **Equation (4.2)** the kinetic energy is

$$k = hf - e_0 = 4.1357 \times 10^{-15} \times \frac{3 \times 10^8}{620 \times 10^{-9}} - 1.6 = 0.4 \qquad \text{[eV]}$$

This kinetic energy is rather low because the red light is close to the longest wavelength to which the photoelectric device responds.

4.4.2.2 Quantum Effects: The Photoconducting Effect

Many modern sensors are based on quantum effects in the solid state, and particularly in semiconductors. Although some electrons may still leave the surface based on the photoelectric effect, when a semiconductor material is subjected to photons they can transfer the energy to electrons. If this energy is sufficiently high, the electrons become mobile, resulting in an increase in the conductivity of the material and, as a result, in an increase in current through the material. This current or its effects become a measure of the radiation intensity (visible light, UV radiation, and, to a lesser extent, IR radiation) that strikes the material. The model for this effect is shown in **Figure 4.2a**. Electrons are normally in the valence band—they are bound to lattice sites within the crystal (i.e., bound to the atoms that make up the crystal) and have specific densities and momentum. Valent electrons are those that are bound to an individual atom. Covalent electrons are also bound, but are shared between two neighboring atoms in the crystal. An electron can only move into the conduction band if its energy is larger than the energy gap (bandgap energy, W_{bg}) specific to the material and if the momentum of the site in the conducting band is the same as the momentum of the electron in the valence band (law of conservation of momentum). This energy may be supplied thermally, but here we are interested in energy absorbed from photons. If the radiation is of sufficiently high frequency (sufficiently energetic photons), valence or covalence electrons may be released from their sites and moved across the bandgap into the conduction band (**Figure 4.2a**).

There are two mechanisms for this transition to occur. In direct bandgap materials, the momentum in the top of the valence band and in the bottom of the conduction band are the same and an electron can transit without the need for a change in momentum, provided it acquires sufficient energy from the photon interaction. In indirect bandgap materials, the electron must interact with the crystal lattice to either gain or lose momentum before it can occupy a site in the conduction band. This process is characterized by a lattice vibration called a phonon and is a less efficient process than that in direct bandgap materials. When in the conduction band, electrons are mobile and free to move as a current. When electrons leave their sites, they leave behind a "hole," which is simply a positive charge carrier. This hole may be taken by a neighboring electron with little additional energy (unlike the original electron released by the photon; see **Figure 4.2b**) and therefore, in semiconductors, the current is due to the net concentrations of electrons and holes. The release of electrons is manifested as a change in the

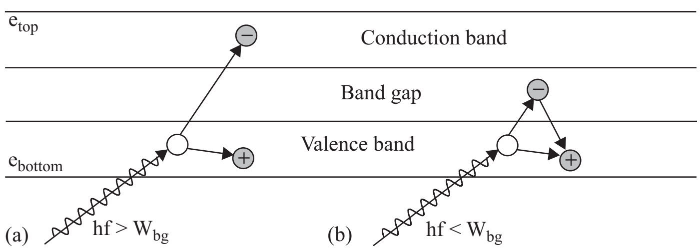


FIGURE 4.2 ■ A model of the photoconductive effect. (a) The photon energy is sufficiently high to move an electron across the bandgap, leaving behind a hole. (b) The photon energy is too low, resulting in recombination of the electron and the hole.

concentration of electrons in the conduction band and of holes in the valence band. The conductivity of the medium is due to the concentrations of both carriers and their mobilities:

$$\sigma = e\left(\mu_e n + \mu_p p\right) \quad [\mathrm{S/m}], \tag{4.4}$$

where σ is conductivity, μ_e and μ_p are the mobilities (in $m^2/V{\cdot}s$ or, often, $cm^2/V{\cdot}s$) of electrons and holes, respectively, and n and p are the concentrations (particles/m^3 or particles/cm^3) of electrons and holes. This change in conductivity or the resulting change in current is then the basic measure of the radiation intensity in photoconducting sensors.

The effect just described is called the **photoconducting effect** and is most common in semiconductors because the bandgaps are relatively small. It exists in insulators as well, but there the bandgaps are very high and therefore it is difficult to release electrons except at very high energies. In conductors the valence and conduction bands overlap (there is no bandgap). Most electrons are free to move, indicating that photons will have minimal or no effect on the conductivity of the medium. Therefore semiconductors are the obvious choice for sensors based on the photoconducting effect, whereas conductors will most often be used in sensors based on the photoelectric effect.

From **Table 4.3**, it is clear that some semiconductors are better suited for low-frequency radiation whereas others are better at high-frequency radiation. The lower the bandgap, the more effective the semiconductor will be at detection at low frequencies (long wavelength, hence lower photon energies). The longest wavelength specified for the material is called the **maximum useful wavelength**, above which the effect is negligible. For example, InSb (indium antimony) has a maximum wavelength of 5.5 μm,

TABLE 4.3 ■ Bandgap energies, longest wavelength, and working temperatures for selected semiconductors

Material	Band gap [eV]	Longest wavelength λ_{max} [μm]	Working temperature [K]
ZnS	3.6	0.35	300
CdS	2.41	0.52	300
CdSe	1.8	0.69	300
CdTe	1.5	0.83	300
Si	1.2	1.2	300
GaAs	1.42	0.874	300
Ge	0.67	1.8	300
PbS	0.37	3.35	
InAs	0.35	3.5	77
PbTe	0.3	4.13	
PbSe	0.27	4.58	
InSb	0.18	6.5	77
Ge:Cu		30	18
Hg/CdTe		8–14	77
Pb/SnTe		8–14	77
InP	1.35	0.95	300
GaP	2.26	0.55	300

Note: Properties of semiconductors vary with doping and other impurities. The values shown should be viewed as representative only.

making it useful in the near IR range. Its bandgap is very low, which also makes it very sensitive. However, that also means that electrons can be easily released by thermal sources, and in fact, the material may be totally useless for sensing at room temperatures (300 K), because at that temperature most electrons will be in the conduction band and these available conduction carriers serve as a thermal background noise for the photon-generated carriers. For this reason, it is often necessary to cool these long-wavelength sensors to make them useful by reducing the thermal noise. The third column in **Table 4.3** shows the (highest) working temperature of the material.

4.4.2.3 Spectral Sensitivity

Each semiconducting material has a range of the spectrum in which it is sensitive, given as a function of frequency or wavelength. The upper range (longest wavelength or minimum energy) is defined by the bandgap (about 1200 nm in **Figure 4.3**). Above the bandgap, the response of a material (i.e., the concentration of conduction electrons in the conduction band due to photon interaction) increases steadily to a maximum and then decreases, as shown schematically in **Figure 4.3**. The reason for the increase and decrease in response is that electron density and momentum are highest in the middle of the valence band and taper off to zero at its boundaries. Because of the law of conservation of momentum, an electron can only transit into the conduction band to a site of like momentum, and the probability of this first increases with an increase in energy (decrease in wavelength), but after most of the electrons in the middle of the valence band have been displaced, the probability of electron transiting to sites of ever-increasing momentum decreases until, at an electron energy equal to the difference between the top of the conduction band and the bottom of the valence band ($e_{top}-e_{bottom}$ in **Figure 4.2**), this probability goes to zero. In **Figure 4.3** that occurs at approximately 650 nm.

4.4.2.4 Tunneling Effect

Another important quantum effect is the **tunneling effect** in semiconductor devices. A simple explanation of this curious effect is that although carriers may not have sufficient energy to go "over" the gap, they can tunnel "through" the gap. Although this explanation is shaky at best, the tunneling effect is real, is a direct consequence of quantum mechanics, and is fully predicted by the Schrödinger equation. The tunneling effect explains behavior on the microscopic level that cannot be explained through classical physics but which, nevertheless, manifests itself on the macroscopic level. Semiconductor devices based on this effect, particularly tunnel diodes, are common, and the effect is used extensively in optical sensors.

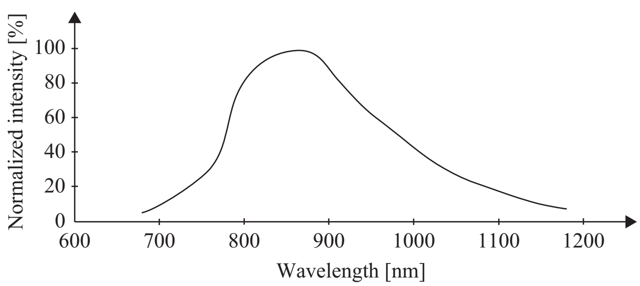


FIGURE 4.3 ■ Spectral sensitivity of a semiconductor.

4.5 | QUANTUM-BASED OPTICAL SENSORS

Optical sensors are divided into two broad classes: **quantum sensors** (or detectors) and **thermal sensors** (or detectors). (Optical sensors are most often called detectors.) A quantum optical sensor is any sensor based on the quantum effects described above and include photoelectric and photoconductive sensors as well as photodiodes and phototransistors (variations of the photoconductive sensor). Thermal optical sensors are mostly encountered in the IR region (and particularly in the low IR region) and come in many variations, including **passive infrared (PIR)** sensors, **active far infrared (AFIR)** sensor bolometers, and others, as we shall see shortly.

4.5.1 Photoconducting Sensors

Photoconducting sensors, or as they are sometimes called, **photoresistive sensors** or **photoresistive cells**, are possibly the simplest optical sensors. They are made from a semiconducting material connected to two conducting electrodes and are exposed to light through a transparent window. A schematic view of the sensor is shown in **Figure 4.4a**. Materials used for these sensors are cadmium sulfide (CdS), cadmium selenide (CdSe), lead sulfide (PbS), indium antimonide (InSb), and others, depending on the range of wavelengths for which the sensor is designed to operate and other requisite properties of the sensor. Of these, CdS is the most common material.

In terms of construction, the electrodes are typically set on top of the photoconductive layer, which in turn is placed on top of a substrate layer. The electrodes may be very simple (**Figure 4.4a**) or may resemble a meandering or comblike shape (**Figure 4.4b**), depending on requirements. In either case, the area exposed between the electrodes is the sensitive area. **Figure 4.5** shows a few sensors of various sizes and construction. The photoconductor is an active sensor that must be connected to a source. The current through or the voltage on the sensors is taken as the output, but what changes with light intensity is the conductivity of the semiconductor and hence its resistance.

The conductivity of the device, given in **Equation (4.4)**, results from the charge of electrons e, the mobilities of electrons and holes (μ_e and μ_p), and the concentrations of electrons n and holes p from whatever source. In the absence of light, the material exhibits what is called dark conductivity, which in turn results in a dark current. Depending on the construction and materials, the resistance of the device may be very high (a few megaohms) or may be in the range of a few kilo-ohms. When the sensor is

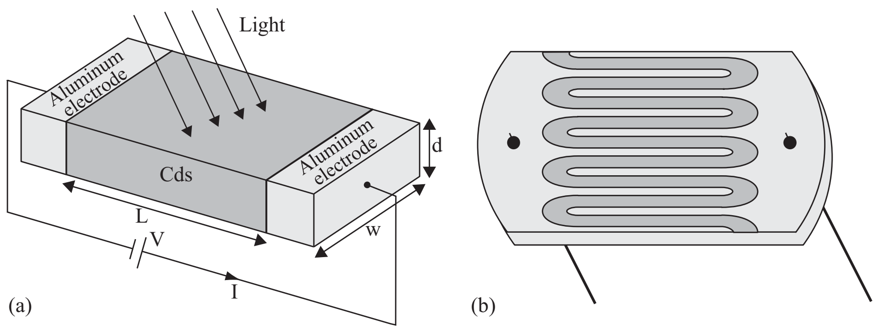


FIGURE 4.4 ■ Structure of a photoconductive sensor. (a) Simple electrodes. (b) Schematic of a sensor showing the connections.

FIGURE 4.5 ■ Examples of photoconductive sensors. The sensor on the right has simple electrodes. The others have comblike electrodes.

illuminated, its conductivity changes (the conductivity increases and hence the resistance decreases) depending on the change in carrier concentrations (excess carrier concentrations). This change in conductivity is

$$\Delta\sigma = e\left(\mu_e \Delta n + \mu_p \Delta p\right) \quad [\mathrm{S/m}], \tag{4.5}$$

where Δn and Δp are the excess carrier concentrations generated by the radiation (light). The carriers are generated by the radiation at a certain generation rate (the number of electrons or holes per second per unit volume), but they also recombine at a set recombination rate. The generation and recombination rates depend on a variety of properties, including the absorption coefficient of the material, dimensions, incident power density (of the radiation) wavelength, and the carrier lifetime (the lifetime of carriers is the time it takes for excess carriers to decay—recombine). Both generation and recombination exist simultaneously, and under a given illumination a steady state is obtained when these are equal. Under this condition, the change in conductivity may be written as

$$\Delta\sigma = ef\left(\mu_n \tau_n + \mu_p \tau_p\right) \quad [\mathrm{S/m}], \tag{4.6}$$

where τ_n and τ_p are the lifetimes of electrons and holes, respectively, and f is the number of carriers generated per second per unit volume. These properties are material dependent and are generally known, although they are temperature as well as concentration dependent. Although carrier generation is in pairs, if the preexisting carrier density of one type dominates, the excess carrier density of the second type will be negligible with respect to the density of the dominant carrier. If electrons dominate, the material is said to be an n-type semiconductor, whereas if the dominant carriers are holes, the semiconductor is said to be p-type. In each of these, the concentration of the opposite type is negligible and the change in conductivity is due to the dominant carrier.

An important property of a photoresistor is its sensitivity to radiation (sometimes called its efficiency). Sensitivity, also called gain, is given as

$$G = \frac{V}{L}\left(\mu_n \tau_n + \mu_p \tau_p\right) \quad [\mathrm{V/V}], \tag{4.7}$$

where L is the length of the sensor (distance between electrodes) and V is the voltage across the sensor. Note that the units in **Equation (4.7)** are volts per volt, hence this is a dimensionless quantity. Sensitivity gives the ratio of carriers generated per photon of the input radiation. To increase sensitivity, one should select materials with high carrier lifetimes, but one must also keep the length of the photoresistor as small as possible. The latter is typically achieved through the meander construction shown in **Figure 4.4b** (see also **Figure 4.5**). The meander shape ensures the distance between two electrodes is reduced for a given exposure area. It also reduces the resistance of the sensor that, referring to **Figure 4.4a**, is given by

$$R = \frac{L}{\sigma wd} \quad [\Omega], \tag{4.8}$$

where wd is the cross-sectional area of the device and σ is its conductivity.

The excess carrier density depends on the power absorbed by the photoconductor. Given a radiation power density P [W/m^2] incident on the top surface of the photoconductor in **Figure 4.4a**, and assuming that a fraction T of this power penetrates into the photoconductor (the rest is reflected off the surface), the power entering the device is $PTS = PTwL$ [W]. This is, by definition, the energy per unit time absorbed in the device. Since the photons have an energy hf, we can write the total number of excess carrier pairs released per unit time as

$$\Delta N = \eta \frac{PTwL}{hf} \quad [\text{carriers/s}], \tag{4.9}$$

where η is the quantum efficiency of the material (a known, given property, dependent on the material used). The latter indicates how efficient the material is at converting photon energy into carriers and clearly indicates that not all photons participate in the process. Assuming carrier generation is uniform throughout the volume of the photoconductor (an assumption only valid for thin photoconductors), we can calculate the rate at which carriers are generated per unit volume per second:

$$\Delta n = \eta \frac{PTwL}{hfwLd} = \eta \frac{PT}{hfd} \quad [\text{carriers/m}^3\text{s}]. \tag{4.10}$$

As mentioned before, the recombination rate influences the net excess carrier density. The carrier density (concentration) is then obtained by multiplying the rate of generation by the lifetime of the carriers, τ:

$$\Delta n = \eta \frac{PT\tau}{hfd} \quad [\text{carriers/m}^3]. \tag{4.11}$$

Both majority and minority excess carriers generated by light have the same densities.

Some of the various terms in **Equation (4.11)** are not necessarily constant with concentrations and some may only be estimates as well as being temperature dependent. However, the equation shows the link between light intensity and excess carrier concentration and hence the dependence of conductivity on light intensity.

Other parameters to consider are the response time of the sensor, its dark resistance (which depends on doping), the range of resistance for the span of the sensor, and the spectral response of the sensor (i.e., the portion of the spectrum in which the sensor is usable). These properties depend on the semiconductor used as well as on the manufacturing processes used to produce the sensor.

Noise in photoconducting sensors is another important factor. Much of the noise is thermally induced and becomes worse at longer wavelengths. Hence many IR sensors must be cooled for proper operation. Another source of noise is the fluctuations in the rates of generation and recombination of the carriers. This noise is particularly important at shorter wavelengths.

From a sensor production point of view, photoresistive sensors are made either as a single crystal semiconductor, by deposition of the material on a substrate, or by sintering (essentially an amorphous semiconductor made of compressed, powdered material sintered at high temperatures to form the photoconductive layer). Usually sensors made by deposition are the least expensive, whereas single crystal sensors are the most expensive, but with better properties. A particular method may be chosen based on requirements. For example, large surface area sensors may need to be made by sintering because large single crystals are both difficult to make and more expensive.

EXAMPLE 4.3 Properties of a photoresistor

Properties of many semiconductors are determined experimentally primarily because of the variability of constituents of the semiconductors and the difficulties of obtaining reliable data. Nevertheless, especially for cadmium sulfide (CdS) sensors, some reliable data are available. To see what the properties of a photoresistor are, consider a simple CdS structure as in **Figure 4.4a** of length 4 mm, width 1 mm, and thickness 0.1 mm. The mobility of electrons in CdS is approximately 210 cm^2/V·s and that of holes is 20 cm^2/V·s. The dark concentration of carriers is approximately 10^{16} carriers/cm^3 (for both electrons and holes). At a light density of 1 W/m^2 the carrier density increases by 11%:

a. Calculate the conductivity of the material and the resistance of the sensor under dark conditions and under the given illumination.
b. Assuming a rate of carrier generation due to light of 10^{15} carriers/cm^3, estimate the sensitivity of the sensor to radiation at a wavelength of 300 nm.

Solution:

a. The conductivity is calculated directly from **Equation (4.4)**:

$$\sigma = e\left(\mu_e n + \mu_p p\right) = 1.602 \times 10^{-19} \times \left(210 \times 10^{16} + 20 \times 10^{16}\right) = 0.36846\ [\mathrm{S/cm}].$$

Because of the units of mobility and carrier density, the result is in siemens per centimeter. By multiplying this by 100 (1 m = 100 cm), we get $\sigma = 36.85$ S/m.

Under light conditions, the carrier density increases by a factor of 1.11 and we get:

$$\sigma = e\left(\mu_e n + \mu_p p\right) = 1.602 \times 10^{-19} \times \left(210 \times 10^{16} \times 1.11 + 20 \times 10^{16} \times 1.11\right) = 0.409\ [\mathrm{S/cm}].$$

The conductivity under light conditions is $\sigma = 40.9$ S/m.

The resistance is found from **Equation (4.9)**:

$$R = \frac{L}{\sigma WH} = \frac{0.004}{36.85 \times 0.001 \times 0.0001} = 1085.5\ [\Omega].$$

$$R = \frac{L}{\sigma WH} = \frac{0.004}{40.9 \times 0.001 \times 0.0001} = 978.0\ [\Omega].$$

Note that the resistance is directly proportional to the increase in carrier density but the increase in carrier density is not linear with illumination. For this reason the resistance decreases rather quickly initially but then levels off, since at high illumination levels there are fewer and fewer carriers available to be released into the conduction band.

b. The sensitivity of the sensor could be calculated from **Equation (4.7)** directly if we had information on the lifetimes of electrons and holes. In their absence we write from **Eqs. (4.5) and (4.6)**:

$$e\left(\mu_e \Delta n + \mu_p \Delta p\right) = ef\left(\mu_n \tau_n + \mu_p \tau_p\right) \rightarrow \left(\mu_n \tau_n + \mu_p \tau_p\right) = \frac{\left(\mu_e \Delta n + \mu_p \Delta p\right)}{f}.$$

Thus we can rewrite **Equation (4.7)** as

$$G = \frac{V}{L^2}\left(\frac{\mu_e \Delta n + \mu_p \Delta p}{f}\right) = \frac{V}{(0.004)^2}\left(\frac{210 \times 10^{-4} \times 1.0 \times 10^{15} + 20 \times 10^{-4} \times 1.0 \times 10^{15}}{10^{15}}\right)$$

$$= 1437/\mathrm{V}.$$

Note that we have converted the units of mobility to $m^2/V \cdot s$ but the units of carrier density do not need to be converted, as they appear in the numerator and denominator similarly, the frequency was writtern as $f = c/\lambda$, with c the speed of light and λ the given wavelength. This gives a sensitivity of 1437/volt, that is, for every 1 V potential difference between the electrodes, a photon generates 1437 carriers. This is a very large sensitivity, typical of CdS sensors.

The most common materials for inexpensive sensors are CdS and CdSe. These offer high sensitivities (on the order of 10^3–10^4), but at a reduced response time, typically about 50 ms. Construction is by deposition and electrodes are then deposited to create the typical comblike shape seen in **Figures 4.4b** and **4.5**, which provides a short distance between the electrodes and a large sensing area. CdS and CdSe can also be sintered. The spectral response of these sensors covers the visible range, although CdS tends to respond better at shorter wavelengths (violet) while CdSe responds better at longer wavelengths (red). Materials can be combined to tailor specific responses. The use of PbS, which is typically deposited as a thin film, shifts the response into the IR region (1000–3500 nm) and improves response to less than 200 μs, but as is typical of IR sensors, at an increase in thermal noise and hence the need for cooling. Examples of single crystal sensors are those made from indium antimonide (InSb). A sensor of this type can operate down to about 7000 nm and can have a response time of less than 50 ns but must be cooled to operate at the longer wavelengths, typically to 77 K (by liquid nitrogen). For specialized application in the IR region, and especially in the far IR, mercury cadmium telluride (HgCdTe) and germanium boronide (GeB) materials may be used. These, especially GeB, can extend operation down to about 0.1 mm if cooled to 4 K (by liquid helium).

In general, cooling of a sensor made of any material extends its spectral response into longer wavelengths, but often slows its response. On the other hand, it increases sensitivity and reduces thermal noise. Many of the far IR applications are military or space applications. These specialized sensors must be made of single crystals and must be housed in a package that is compatible with the low temperature requirements.

4.5.2 Photodiodes

If the junction of a semiconducting diode is exposed to light radiation, the generation of excess carriers due to photons adds to the existing charges in the conduction band

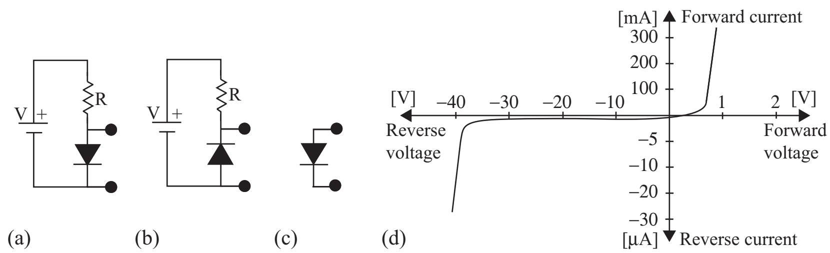


FIGURE 4.6 ■ The semiconducting (*p-n*) junction. (a) Forward biased. (b) Reverse biased. (c) Unbiased. (d) The *I-V* characteristics of the junction.

exactly in the same fashion as for a pure semiconductor. The diode itself may be reverse biased (**Figure 4.6a**), forward biased (**Figure 4.6b**), or unbiased (**Figure 4.6c**). **Figure 4.6d** shows the current–voltage (*I-V*) characteristics of the diode. Of the three configurations in **Figure 4.6**, the forward-biased mode is not useful as a photosensor because in this mode the normal current (not due to photons) is large in comparison to the current generated by photons. In the reverse-biased mode, the diode carries a minute current (i.e., a "dark" current) and the increase in current due to photons is large in comparison. In this mode the diode operates in a manner similar to the photoconducting sensor and is therefore called the **photoconducting mode** of the diode. If the diode is not biased it operates as a sensor in the **photovoltaic mode** (**Figure 4.6c**).

The equivalent circuit of a diode in the photoconductive mode (**Figure 4.6b**) is shown in **Figure 4.7a**. In addition to the current that would exist in the ideal diode (I_d) there is also a leakage current (I_0) defined by the "dark" resistance R_0 and a current through the capacitance (I_c) of the junction. The series resistance R_s is due to conductors connecting the diode. The photons release electrons from the valence band either on the *p* or *n* side of the junction. These electrons and the resulting holes flow toward the respective polarities (electrons toward the positive pole, holes toward the negative pole) generating a current, which in the absence of a bias current in the diode constitutes the only current (the diode is reverse biased). In practice there will be a small leakage current, shown in the equivalent circuit as I_0. The attraction of electrons by the positive pole will tend to accelerate them and in the process they can collide with other electrons and release them across the bandgap, especially if the reverse voltage across the diode is high. This is called an **avalanche effect** and results in multiplication of the carriers available. Sensors that operate in this mode are called **photomultiplier sensors**.

In any diode the current in the forward-biased mode is

$$I_d = I_0\left(e^{eV_d/nKT} - 1\right) \qquad [\text{A}] \tag{4.12}$$

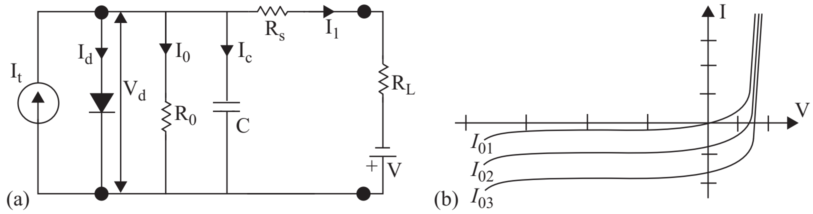


FIGURE 4.7 ■ A photodiode connected in the photoconductive mode (reverse biased). (a) Equivalent circuit. (b) *I-V* characteristics.

where I_0 is the leakage (dark) current, e is the charge of the electron, V_d is the voltage across the junction, also called the voltage barrier or built-in potential, k is Boltzmann's constant ($k = 1.3806488 \times 10^{-23}$ m^{2}kg/s^2/K or J/K), T is the absolute temperature [K], and n is an efficiency constant between 1 and 2. In practical cases it is equal to 1 (but see **Section 3.4**, where it was equal to 2). This relation clearly indicates the dependence of the diode current on temperature as well as on the bias voltage. However, in the reverse mode, only the current I_0 may flow and for most practical purposes it is very small, often negligible.

The current produced by photons is given as

$$I_p = \frac{\eta PAe}{hf} \quad [\mathrm{A}], \tag{4.13}$$

where P is the radiation (light) power density [W/m^2], f is the frequency, and h is Planck's constant. All other terms are constants associated with the diode or the semiconductor used. η is the quantum absorption efficiency, A is the exposed area of the diode (ηPA is the power absorbed by the junction). The total current available external to the diode is (using $n = 1$)

$$I_l = I_d - I_p = I_0\left(e^{eV_d/KT} - 1\right) - \frac{\eta PAe}{hf} \qquad [\mathrm{A}] \tag{4.14}$$

This is the current measured for a photodiode sensor under forward or backward biasing conditions (depending on the voltage V_d across the junction). Under reverse-bias conditions, the first term may be neglected altogether since I_0 is small (on the order of 10 nA) and V_d is negative. As a first approximation, especially at low temperatures, one obtains a simple relation for the photodiode:

$$I_l \approx \frac{\eta PAe}{hf} \quad [\mathrm{A}]. \tag{4.15}$$

When measured, this current gives a direct reading of the power absorbed by the diode, but as can be seen, it is not constant since the relation depends on frequency and the power absorbed itself is frequency dependent except for monochromatic radiation. As the input power increases, the characteristic curve of the diode changes, as shown in **Figure 4.7b**, resulting in an increase in the reverse current, as expected.

Any diode can serve as a photodiode, provided that the *n* region, *p* region, or *p-n* junction are exposed to radiation. However, specific changes in materials and construction have been made to common diodes to improve one or more of their photoconducting properties (usually the dark resistance and response time). Taking as an example the planar diffusion type of diode shown in **Figure 4.8**, it consists of *p* and *n* layers and two contacts. The region immediately below the *p* layer is the so-called depletion region, which is characterized by an almost total absence of carriers. This is essentially a regular diode. To increase dark resistance (lower dark current), the *p* layer may be covered with a thin layer of silicon dioxide (SiO_2) (**Figure 4.8a**). The addition of an intrinsic layer of the semiconductor between the *p* and *n* layers produces the so-called PIN photodiode, which, because of the high resistance of the intrinsic layer, has lower

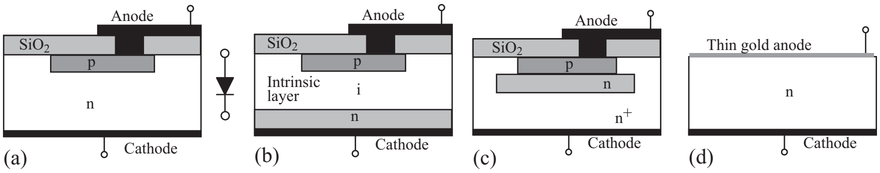


FIGURE 4.8 ■ Various structures of photodiodes. (a) Common planar structure. (b) PIN diode. (c) pnn^+ structure. (d) Schottky diode.

dark current and lower junction capacitance (and hence better time response) (**Figure 4.8b**). The exact opposite is true in the *pnn*+ construction in which a thin, highly conductive layer is placed at the bottom of the diode. This reduces the resistance of the diode and improves low-wavelength sensitivity (**Figure 4.8c**). Another way of altering the response of a diode is through the use of a Schottky junction. In this diode, the junction is formed by use of a thin layer of sputtered conducting material (gold) on an *n* layer (the Schottky junction is a metal–semiconductor junction) (**Figure 4.8d**). This produces a diode with a very thin outer layer (metal) above the *n* layer, improving its long wavelength (IR) response. As mentioned before, a diode with high reverse bias may operate in avalanche mode, increasing the current and providing a gain or amplification (**photomultiplier diode**). The main requirement needed to obtain avalanche is the establishment of a high reverse electric field across the junction (on the order of 10^7 V/m or higher) to provide sufficient acceleration of electrons. In addition, low noise is essential. Avalanche photodiodes are available for high-sensitivity, low light level applications.

Photodiodes are available in various packages, including surface mount, plastic, and small can packages. **Figure 4.9a** shows one type of diode used as a detector for reflected laser light in CD players. They are also available in linear arrays of photodiodes, such as in **Figure 4.9b**, which shows a (512 element) linear array used as the sensor for a scanner. They are available for IR as well as for the visible range and some extend the range into the UV and even the X-ray range. Many photodiodes have a simple lens to increase the power density at the junction.

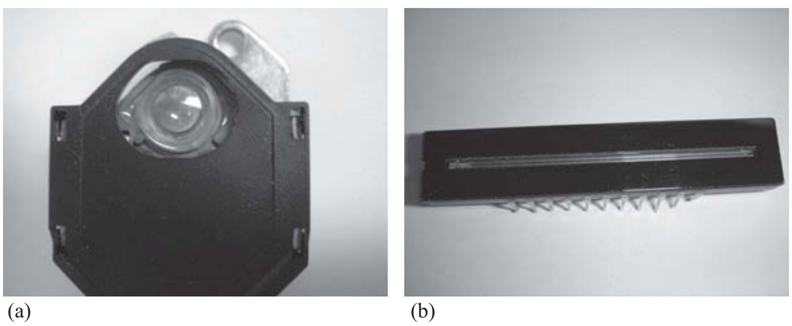


FIGURE 4.9 ■ (a) A photodiode used as a sensor in a CD player shown installed in its holder. (b) A photodiode linear array (512 photodiodes) in a single integrated circuit used as the sensor element in a scanner. The top cover is glass and light is allowed in through the transparent slit.

EXAMPLE 4.4 Photodiode as a detector for fiber-optic communication

A digital communication link uses a red laser operating at 800 nm with an output of 10 mW. The optical link is 16 km long and is made of an optical fiber with an attenuation of 2.4 dB/km. At the receiving end of the link, a photodiode detects the pulses and the output is measured across a 1 MΩ resistor, as shown in **Figure 4.10**. Assuming the transmitter transmits a series of pulses and there are no losses on either end, that is, all power produced by the laser enters the optical fiber and all power at the diode is absorbed by the diode, calculate the amplitude of the received pulses. Assume that the dark current (leakage current) of the diode is 10 nA and the system operates at 25°C. How does the amplitude change if the temperature rises to 50°C?

Solution: We calculate the diode current using **Equation (4.12)** followed by the photon current using **Equation (4.13)**, or alternatively we can calculate the total current using **Equation (4.14)**. However, to do so we need the radiation power density P. That is calculated using the incident power density and the attenuation along the optical fiber as follows:

The laser produces 10 mW, but to calculate the power entering the diode we need to take into account the losses. To do so we first calculate the input power, P, in decibels:

$$P = 10 \times 10^{-3}\ \mathrm{W} \rightarrow P = 10\log_{10}\left(10 \times 10^{-3}\right) = -20\ \mathrm{dB}.$$

The total attenuation along the line is $2.4 \times 16 = 38.4$ dB. Therefore the power in decibels at the end of the line is

$$P = -20 - 38.4 = -58.4\ \mathrm{dB}.$$

Now we convert this back to power. To do so we write

$$10\log_{10} P_0 = -58.4 \rightarrow \log_{10} P_0 = -5.84 \rightarrow P_0 = 10^{-6.84} = 1.445 \times 10^{-6}\ \mathrm{W}.$$

Now we can use **Equation (4.14)**, but we first note that the term ηPA is the total power received by the diode. That, according to the problem statement, equals P_0. Thus we write

$$\begin{aligned} I_l &= I_0\left(e^{eV_d/kT} - 1\right) - \frac{P_0 e}{hf} = 10 \times 10^{-9}\left(e^{1.61\times10^{-19}\times(-12)/1.3806488\times10^{-23}\times298} - 1\right) \\ &\quad - \frac{1.445 \times 10^{-6} \times 1.61 \times 10^{-19}}{6.6262 \times 10^{-34} \times 3.75 \times 10^{14}} \\ &= -10 \times 10^{-9} - 936.3 \times 10^{-9}\ \mathrm{A} \end{aligned}.$$

Clearly the current due to temperature is negligible. Now we can calculate the output voltage across the resistor. When the laser beam is off, the current through the photodiode is 10 nA and the output voltage is

$$V_0 = 10 \times 10^{-9} \times 1 \times 10^{6} = 10\ \mathrm{mV}.$$

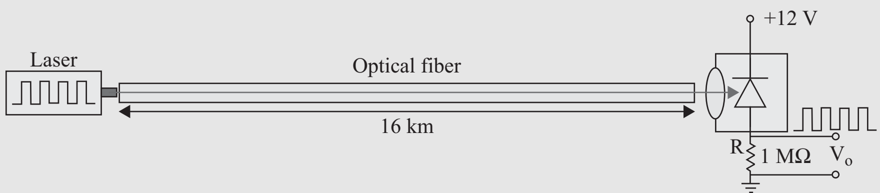


FIGURE 4.10 ■ An optical fiber communication link with the source (laser) and a photodiode used as a detector.

When the light beam is on, the current increases to 946.3 nA and the output voltage is

$$V_0 = 946.3 \times 10^{-9} \times 1 \times 10^6 = 946.3 \text{ mV}.$$

That is, the pulses will result in a voltage that changes from 10 mV for level "0" to 0.946 V for level "1" of the pulse. This may not be sufficient for interfacing and may require amplification.

At 50°C the solution is essentially the same since the component of the current that depends on temperature is negligible in this case. However, this is not a general conclusion.

4.5.3 Photovoltaic Diodes

Photodiodes may also operate in photovoltaic mode as shown in **Figure 4.11**. In this mode the diode is viewed as a generator and requires no biasing. The best known structure for photovoltaic diodes is the solar cell, which is a photodiode with a particularly large exposed area. All photodiodes can operate in this mode, but as a rule, the larger the surface area, the larger the junction capacitance. This capacitance is the main reason for the reduced time response of photovoltaic cells. In most other respects photodiodes operating in photovoltaic mode have the same properties as photodiodes operating in photoconducting mode. There are differences as well. For example, the avalanche effect cannot exist in this mode since there is no bias. **Figure 4.11** shows the equivalent circuit for a photovoltaic cell and **Figure 4.12** shows two small photovoltaic (solar) cells.

Although typically used in photovoltaic arrays for solar power generation as well as smaller arrays used to power small appliances (such as calculators), the photovoltaic diode also makes an exceedingly simple light sensor that needs little more than a voltmeter to measure light power density or light intensity.

Although the photovoltaic diode operates without a bias, under normal operation a voltage develops across the junction and the total current is described by **Equation (4.14)**, where the first term is the normal diode current and the second is the photocurrent. There are two important points in the operation of the diode. The first is the short-circuit current. If the diode is short-circuited, the voltage across the diode is zero and the only current that may exist is the photocurrent. Thus

$$I_{sc} = -I_p = -\frac{\eta P A e}{hf} \quad [\text{A}]. \tag{4.16}$$

The second term is the open circuit voltage, characterized by the fact that the normal diode current equals the photocurrent. The open circuit voltage, V_{oc}, can then be evaluated from this balance:

$$I_0\left(e^{eV_{oc}/kT} - 1\right) = -\frac{\eta P A e}{hf} = I_p. \tag{4.17}$$

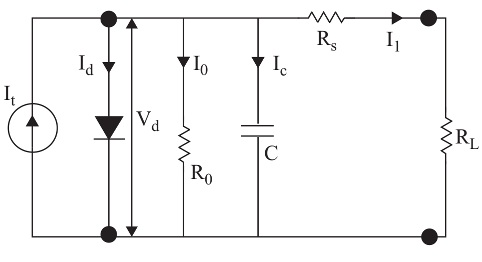


FIGURE 4.11 ■ Photodiode connected in the photovoltaic mode—equivalent circuit. The diode is unbiased.

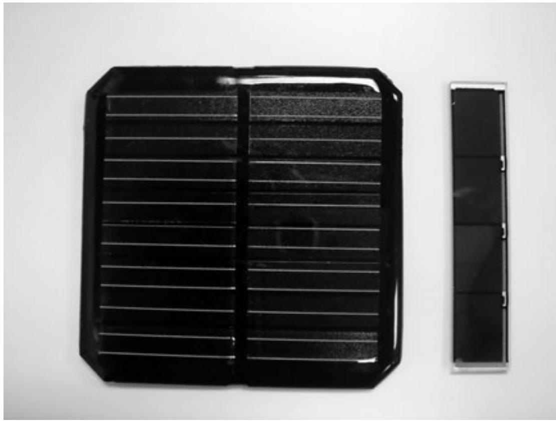


FIGURE 4.12 ■ A particular type of photodiode, the photovoltaic cell or solar cell. Two types are shown. The one on the right is used in a calculator.

Clearly V_{oc} is equal to the built-in potential or potential barrier, which depends on the material, doping, and, through carrier concentration, temperature. The efficiency term in **Equations (4.16)** and **(4.17)** is an overall efficiency of the cell and is a product of the quantum absorption coefficient and the conversion coefficient of the cell.

EXAMPLE 4.5 Properties of a solar cell at low light

To establish the transfer function of a solar cell one must measure the input power or power density and the output voltage or power supplied by the cell. In particular, at low light, the conversion efficiency of the cell is low and the transfer function is very much dependent on the load (R_L in **Figure 4.11**). A solar cell, 11 cm × 14 cm in area, is connected to a 1 kΩ load, exposed to light from an artificial source, and the output voltage is measured. The voltage versus input power density is shown in **Figure 4.13a**. The curve is characteristic, indicating that the cell tends toward a saturation output at some (higher) power density. This particular solar cell only supplies a few milliwatts because of the low illuminance level.

One of the important parameters of solar cells is its conversion efficiency. Defined as the output power divided by the input power, it is usually given in percentages. As with all other

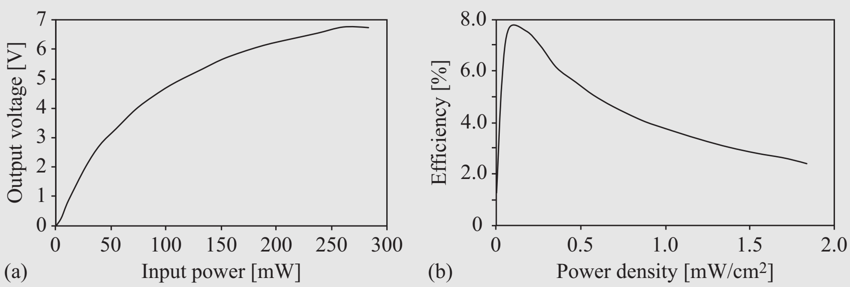


FIGURE 4.13 ■ Characteristics of a solar cell at low light. (a) Output voltage versus input power density. (b) Power conversion efficiency versus input power density.

characteristics, it depends on load and operating point (input power). The efficiency of the cell used here is calculated as follows:

$$e = \frac{P_{out}}{P_{in}} \times 100\% = \frac{V^2/R_L}{P_d \times S} \times 100\%,$$

where P_d is the input power density [mW/cm^2], S is the area of the cell [cm^2] (in this case 11 cm $\times$ 14 cm = 154 cm^2), R_L is the load [kΩ], and V is the output voltage [V]. The efficiency is plotted in **Figure 4.13b**. The efficiency reaches a maximum of 7.78% at 0.174 mW/cm^2 and declines beyond that, indicating the same tendency toward saturation. It should be noted again that the maximum efficiency point depends on the illuminance level and the load. Good solar cells will have efficiencies between 15% and 30%.

EXAMPLE 4.6 Optical position sensor

A position sensor is made as follows: A triangular slit is cut into an opaque material that covers a solar cell to create a triangular area of exposure (**Figure 4.14**). This makes the stationary part of the sensor. A moving part is placed above the stationary part and includes two items. One is a stationary source of light, and between the source and the stationary slit is a thin rectangular opening that allows light to go through that opening only. The opening is t m wide and the source supplies an illuminance of I lux. The position of the slit is the output of the sensor. Since the larger h (distance) is, the larger the light power on the cell, the larger the voltage of the solar cell. Assuming the voltage output to be linear with the power incident on the solar cell following the relation $V = kP$, where P is the incident power and k a constant of the cell, find a relation between the measured voltage of the cell and the position of the slit, h.

Solution: As h increases so does the amount of light reaching the cell because the width of the illuminated slit increases. Illuminance is measured in lux, which has units of watts per square meter. Therefore the power reaching the cell is proportional to the area of the triangular slit covered by the rectangle of width t. We calculate this area, multiply by the illuminance, and find the power P. It should be recalled, however, that 1 lux is (1/683) W/m^2. Then of course the output voltage is immediately available.

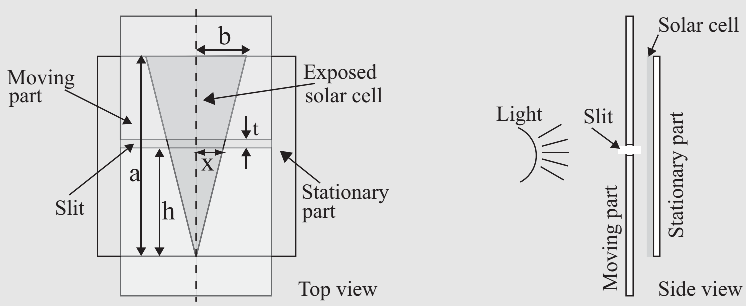


FIGURE 4.14 ■ An optical position sensor. The slit in the top layer indicates the position by exposing a strip width that depends on position.

Assuming the slit is at a height h as shown, the area exposed is

$$S = \frac{(2x' + 2x)t}{2} = (x' + x)t \quad [\mathrm{m}^2].$$

To calculate x and x' we note the following:

$$\frac{b}{a} = \frac{x}{h} = \frac{x'}{h+t}.$$

Therefore

$$x = \frac{b}{a}h, x' = \frac{b}{a}(h+t) \quad [\mathrm{m}].$$

We have

$$S = \left(\frac{b}{a}(h+t) + \frac{b}{a}h\right)t = \frac{bt}{a}(2h+t) \quad [\mathrm{m}^2].$$

The power reaching the cell is

$$P = SI = \frac{btI}{a}(2h+t) \quad [\mathrm{W}],$$

where I is the illuminance [W/m^2]. The output voltage is

$$V = kP = k\frac{btI}{a}(2h+t) \quad [\mathrm{V}].$$

Since the output of the sensor is voltage and the measurand is the distance h, we write this in a more convenient way:

$$V = 2k\frac{btI}{a}h + k\frac{bt^2I}{a} \quad [\mathrm{V}].$$

A linear relation is obtained as expected. Note also that the sensitivity can be increased by increasing b, t, and/or I.

4.5.4 Phototransistors

As an extension of the discussion on photodiodes, the phototransistor can be viewed as two diodes connected back to back, as shown in **Figure 4.15** for an *npn* transistor. With the bias shown, the upper diode (the collector–base junction) is reverse biased while the

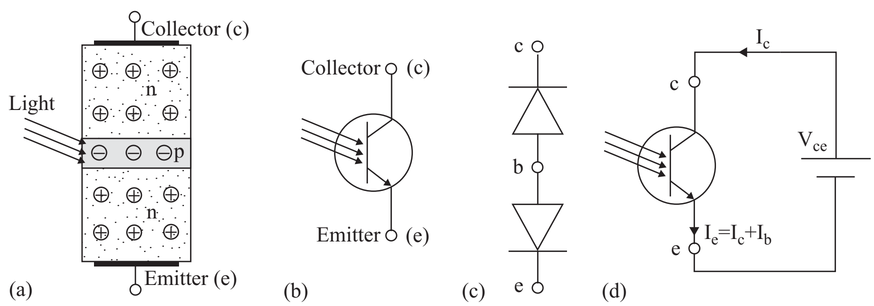


FIGURE 4.15 ■ An *npn* phototransistor. (a) Schematic structure and junctions. (b) The circuit schematic. (c) The two junctions form diodes as shown. (d) The biasing of a phototransistor.

lower (base–emitter) junction is forward biased. In a regular transistor, a current I_b injected into the base is amplified using the following simple relation:

$$I_c = \beta I_b \quad [\mathrm{A}], \tag{4.18}$$

where I_c is the collector current and β is the amplification or gain of the transistor, which depends on a variety of factors, including bias, construction, materials used, doping, etc. The emitter current I_e is

$$I_e = I_b(\beta + 1) \quad [\mathrm{A}]. \tag{4.19}$$

The relations above apply to any transistor. What is unique in a phototransistor is the means of generating the base current. When a transistor is made into a phototransistor, its base connection is usually eliminated and a provision is made for the radiation to reach the collector–base junction. The device operates as a regular transistor with its base current supplied by the photon interaction with the collector–base junction (which is reverse biased). The transistor described here is also called a bipolar junction transistor (BJT). This name distinguishes it from other types of transistors, some of which we will encounter later.

Under dark conditions, the collector current is small and is almost entirely due to leakage currents, designated here as I_0. This causes a dark current in the collector and emitter as

$$I_c = I_0\beta, I_e = I_0(\beta + 1) \quad [\mathrm{A}]. \tag{4.20}$$

When the junction is illuminated, the diode current is the current due to photons obtained in **Equation (4.13)**:

$$I_b = I_p = \frac{\eta PAe}{hf} \quad [\mathrm{A}]. \tag{4.21}$$

The collector and emitter currents are then

$$I_c = I_p\beta = \beta\frac{\eta PAe}{hf}, I_e = I_p(\beta + 1) = (\beta + 1)\frac{\eta PAe}{hf} \quad [\mathrm{A}], \tag{4.22}$$

where the leakage current was neglected in the final relations, as was done for the photodiode. Clearly then the operation of the phototransistor is identical to that of the photodiode except for the amplification, β, provided by the transistor structure. Since β, for even the simplest transistors, is on the order of 100 (and can be much higher), and the amplification is linear in most of the operation range (see **Figure 4.16**), the phototransistor is a very useful device and is commonly used for detection and sensing. The high amplification allows phototransistors to operate at low illumination levels. On the

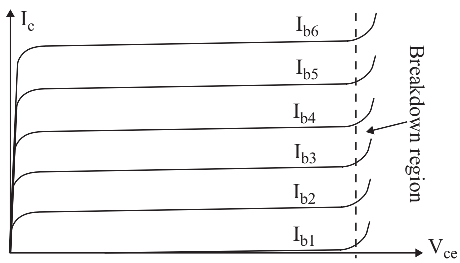


FIGURE 4.16 ■ The *I-V* characteristics of a transistor as a function of base current. In a phototransistor the base current is supplied by photon interaction.

FIGURE 4.17 ■ Phototransistors equipped with lenses.

other hand, thermal noise can be a bigger problem, again because of the amplification. In particular, the base–emitter junction behaves as a regular diode as far as current through it. The latter is given in **Equation (4.12)**, where again, I_0 is the dark current. Although this current is small, the fact that the diode is forward biased, and due to the amplification of the transistor, the effects of temperature are significant.

In many cases a simple lens is also provided to concentrate the light on the junction, which for transistors is very small. A phototransistor equipped with a lens is shown in **Figure 4.17**.

Photoconducting sensors, photodiodes, and phototransistors can sense and measure directly the radiation power they absorb. However, they can easily be used to sense any other quantity or effect that can be made to generate or alter radiation in the range in which the sensor is sensitive. As such, they can be employed to sense position, distance, and dimensions, temperature, and color variations, in counting events, for quality control, and much more.

EXAMPLE 4.7 Sensitivity of a phototransistor

Figure 4.16 shows the *I-V* characteristics of a transistor as a function of base current. However, in a phototransistor, the base current is not measurable. Rather, the current is a function of light power density on the junction. The following is an experimental evaluation of the current in a phototransistor as a function of incident light power density. The table below shows selected values, but all values measured are plotted in **Figure 4.18**. Since the curve is linear for light densities between 0 and about 400 μW/cm^2, the sensitivity of the sensor can be written using any two columns in the table. Taking the first column and the eighth column, we find its sensitivity in the linear range as

$$S = \frac{0.13 - 0.00182}{152 - 2} = \frac{0.12818}{150} = 0.8545\ \mu\text{A}/\mu\text{W}/\text{cm}^2.$$

Light density [μW/cm^2]	2	9.57	20.7	46.2	60.4	83.9	113	152	343	409
Current [μA]	.00182	.00864	.0182	.0409	.0532	.0732	.0978	.13	.28	.324

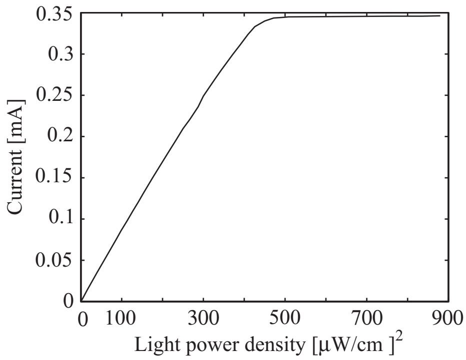


FIGURE 4.18 ■ Collector current in a phototransistor as a function of input light power density. Note the saturation above 400 μW/cm^2.

4.6 | PHOTOELECTRIC SENSORS

Photoelectric sensors, including photomultipliers, are based on the photoelectric effect (also called the photoemissive effect). As described in **Section 4.4.2**, the photoelectric effect relies on photons with energy *hf* impinging on the surface of a material. The radiation is absorbed by giving this energy to electrons and these are emitted from the surface, provided the energy of the photon is higher than the work function of the material. One can say that the collision between the photon and electron releases the electron if the energy exchanged is sufficiently high. This effect has been applied directly to the development of photoelectric sensors (sometimes called photoelectric cells). In fact, this type of optical sensor is one of the oldest optical sensors available.

4.6.1 The Photoelectric Sensor

The principle of the photoelectric sensor is shown in **Figure 4.19**. The photocathode is made of a material with relatively low work function to allow efficient emission of electrons. These electrons are then accelerated toward the photoanode because of the potential difference between the anode and cathode. The current in the circuit is then proportional to radiation intensity. The number of emitted electrons per photon is the quantum efficiency of the sensor and depends to a large extent on the material used for the photocathode (its work function). Many metals may be used for this purpose, but for the most part their efficiencies are low. More often, cesium-based materials are used because they have low work functions and fairly wide spectrum responses down to about 1000 nm (well into the IR region). Their response extends into the UV region as well. In older devices, highly resistant cathodes were made of a metal such

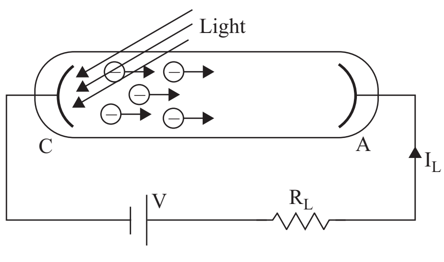


FIGURE 4.19 ■ Photoelectric sensor and biasing circuit.

as tantalum or chromium and coated with alkali compounds (lithium, potassium, sodium, or cesium, or more often, a combination of these; see the periodic table). This provides the low work function necessary. The electrodes are housed in an evacuated tube or in a tube with a noble gas (argon) at low pressure. The presence of gas increases the gain of the sensor (defined as the number of electrons emitted per incoming photon) by internal collisions between emitted electrons and the atoms of the gas through ionization of the gas. Newer devices use so-called negative electron affinity (NEA) surfaces. These are constructed by evaporation of cesium or cesium oxide onto a semiconductor's surface.

Classical photoelectric sensors require relatively high voltages for operation (sometimes a few hundred volts) to supply useful sensing currents. Negative electron afirity devices operate at much lower potentials.

4.6.2 Photomultipliers

Photomultipliers are a development of the classical photoelectric sensor. Whereas in a photoelectric sensor the current is low (the number of electrons emitted is small), photomultipliers, as their names imply, multiply the available current, resulting in sensors that are considerably more sensitive than the simple photoelectric cell. The construction is shown schematically in **Figure 4.20**. It consists of an evacuated tube (or a low-pressure gas-filled tube) made of metal, glass, or metal-coated glass with a window for the incoming radiation. The photocathode and photoanode of the basic photoelectric cell are maintained, but now there is a sequence of intermediate electrodes, as shown in **Figure 4.20a**. The intermediate electrodes are called **dynodes** are made of materials with low work functions, such as beryllium copper (BeCu), and are placed at potential differences with respect to preceding dynodes, as shown in **Figure 4.20b**. The operation is as follows: The incident radiation impinges on the cathode and releases a number of electrons, say, n. These are accelerated toward the first dynode by the potential difference, V_1. These electrons now have sufficient energy to release, say, n_1 electrons for each impinging electron. The number of electrons emitted from the first dynode is $n \times n_1$. These are again accelerated toward the second dynode, and so on, until they finally reach the photoanode. The multiplication effect at each dynode results in a very large number of electrons reaching the photoanode for each photon impinging on the cathode. Assuming there are k dynodes (10–14 is not unusual) and n is the average number of electrons emitted per dynode (secondary electrons), the gain may be written as

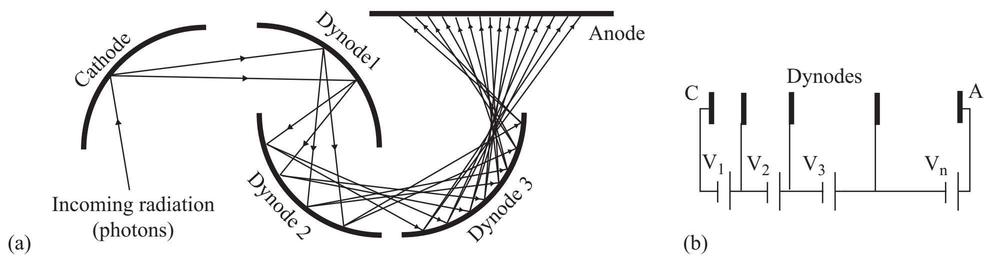


FIGURE 4.20 ■ (a) Basic structure of a photomultiplier. (b) Biasing of the dynodes and photoanode. The typical potential difference between an anode and a cathode is about 600 V, about 60–100 V between each two dynodes.

FIGURE 4.21 ■ A photomultiplier. The light enters through the top surface. The dynodes are the curved surfaces on top.

$$G = n^k. \tag{4.23}$$

This gain is the current amplification of the photomultiplier and depends on the construction, the number of dynodes, and the accelerating interelectrode voltages. Clearly, additional considerations must be employed for maximum performance. First, electrons must be "forced" to transit between electrodes at about the same time to avoid distortions in the signal. To do so the dynodes are often shaped as curved surfaces that also guide the electrons toward the next dynode. Additional grids and slats are added for the same purpose, to decrease transit time and improve quality of the signal, especially when the photomultiplier is used for imaging.

As with all sensors of this type, there are sources of noise, but because of the multiplying effect, noise is particularly important in photomultipliers. Of these, the dark current due to thermal emission, which is both potential and temperature dependent, is the most critical. The dark current in a photomultiplier is given as

$$I_0 = aAT^2 e^{-E_0/kT}, \tag{4.24}$$

where a is a constant depending on the cathode material, generally around 0.5, A is a universal constant equal to 120.173 A/cm^2, T is the absolute temperature, E_0 is the work function of the cathode material [e·V], and k is Boltzmann's constant. In photomultipliers this current is small because the cathode is cold and under these conditions the thermal emission is low. Nevertheless, a dark current between 1 and 100 nA is present because of the high gain of the photomultiplier. In addition, the shot noise due to fluctuations in the current of discrete electrons and multiplication noise due to the statistical spread of electrons limits the sensitivity of the device. A major concern with photomultipliers is their susceptibility to magnetic fields. Since magnetic fields apply a force on moving electrons, they can force electrons out of their normal paths reducing their gain and, more critically, distorting the signal.

Nevertheless, with proper construction (including possible cooling of the sensor to reduce dark current) an exceedingly sensitive device can be made. These sensors are therefore used for very low light applications, such as in night vision systems. For example, a photomultiplier sensor may be placed at the focal point of a telescope to view extremely faint objects in space.

Photomultipliers are part of a broader class of devices called **image intensifiers** that use various methods (including electrostatic and magnetic lenses) to increase the current due to radiation. Because their output is sometimes the image itself, they are sometimes called light-to-light detectors. **Figure 4.21** shows a small photomultiplier.

These devices have many disadvantages, including problems with noise, as discussed above, size, the need for high voltages (in excess of 2000 V for some), as well as cost. For these reasons, except for some applications in night vision, they have been largely displaced by charge-coupled devices (CCDs), which have many of the advantages of photomultipliers while eliminating most of the problems associated with photomultipliers.

EXAMPLE 4.8 **Thermionic noise in a photomultiplier**

A photomultiplier with 10 dynodes has a cathode coated with potassium to increase sensitivity. Calculate the thermally produced dark current at the cathode and at the anode at 25°C, assuming that each incoming photon is energetic enough to release six electrons and that each accelerated electron releases six electrons.

Solution: The work function of potassium is 1.6 eV (see **Table 4.2**). Room temperature is $273.15 + 25 = 298.15°K$. With the Boltzmann constant, $k = 8.62 \times 10^{-5}$ eV/°K, we get

$$I_0 = aAT^2 e^{-E_0/kT} = 0.5 \times 120.173 \times 10^4 \times (298.15)^2 e^{-1.6/8.62\times10^{-5}\times298.15} = 4.9 \times 10^{-17} \text{ A}.$$

This is a mere 4.9×10^{-8} nA. Since each accelerated electron releases six electrons, the gain of the photomultiplier is

$$G = n^k = 6^{10} = 6.05 \times 10^7.$$

The current at the anode due to thermionic emission is

$$I_a = 4.9 \times 10^{-17} \times 6.05 \times 10^7 = 2.96 \times 10^{-9} \text{ A}.$$

This is just under 3 nA.

It is because of these very low dark currents that the photomultiplier is so very useful and has survived into the age of semiconductors.

4.7 | COUPLED CHARGE (CCD) SENSORS AND DETECTORS

Charge coupled devices (CCD) are typically made of a conducting substrate on which a *p*- or *n*-type semiconductor layer is deposited. Above it lies a thin insulating layer made of silicon dioxide to insulate the silicon from a transparent conducting layer above it, as shown in **Figure 4.22a**. This structure is called a metal oxide semiconductor (MOS) and is a simple and inexpensive structure. The conductor (also called a gate) and the substrate form a capacitor. The gate is biased positively with respect to the substrate (for an *n*-type semiconductor). This bias causes a depletion region in the semiconductor and, together with the silicon dioxide layer, makes this structure a very high resistance device. When optical radiation impinges on the device it penetrates through the gate and

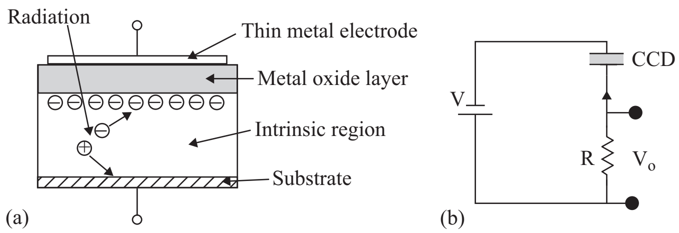


FIGURE 4.22 ■ The basic CCD cell. (a) In forward-biased mode, electrons accumulate below the MOS layer. (b) In reversed-biased mode the charge is sensed by discharging it through an external load.

oxide layer to release electrons in the depletion layer. The charge density released is proportional to the incident radiation intensity. These charges are attracted toward the gate but cannot flow through the oxide layer and are trapped there. There are a number of methods to measure the charge (and hence the radiation intensity that produced it). In its simplest form, one can reverse bias the MOS device to discharge the electrons through a resistor, as shown in **Figure 4.22b**. The current through the resistor is a direct measure of the light intensity on the device. However, the main value of CCDs is in building one-dimensional (linear array) or two-dimensional arrays of MOS devices for the purpose of imaging. In such cases it is not possible to use the method in **Figure 4.22b** directly. Rather, the basic method is to move the charges of each cell to the next in a kind of "musical chairs" sequence by manipulating gate voltages. In this method the transfer is one cell per step and the current in the resistor for each step corresponds to a particular cell. This is shown in **Figure 4.23a** for one row in the two-dimensional array. At the end of this scan, the array can be scanned again to read a new image. The scan for a two-dimensional array is shown schematically in **Figure 4.23b**. The data are moved vertically one row at a time, that is, all cells move their data one row lower, whereas the lowest row moves it into a shift resister. The scan stops and the shift register is moved to the right to obtain the signal for one row (similar to that in **Figure 4.23a**). Then the next row is shifted until the whole array has been scanned. In practice, each cell is equipped with three electrodes, each covering one-third of the cell and the time step described above is made of three pulses or phases. All first electrodes in a row are connected to each other, all second electrodes in a row form a

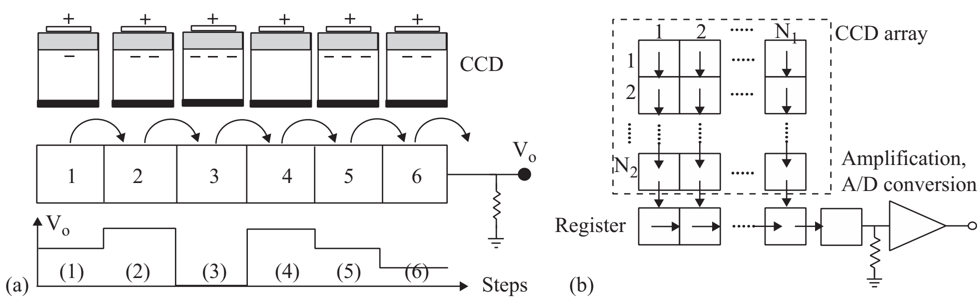


FIGURE 4.23 ■ Method of sensing the charge in a CCD array. (a) The charge is moved in steps to the edge (by manipulating the gate voltages) and discharged through a resistor. (b) Two-dimensional scan of an $N_1 \times N_2$ image.

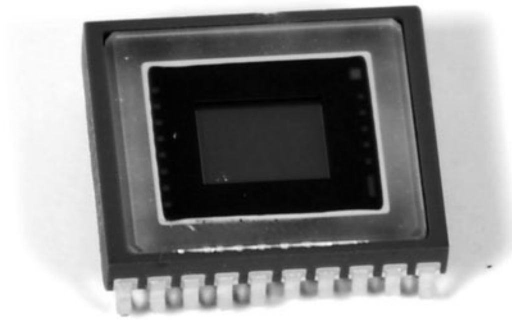


FIGURE 4.24 ■ A CCD used as the imaging element in a small camera.

second phase, and all third electrodes in a row form the third phase. The phases are powered in sequence, moving all charge in each row downward one-third of the cell. After three pulses the charge of each row is transferred to the row below. The signal obtained is typically amplified and digitized and used to produce the image signal, which can then be displayed on a display array such as a TV screen or a liquid crystal display. Of course, there are many variations of this basic process. For example, to sense color, filters may be used to separate colors into their basic components (red-green-blue [RGB] is one method). Each color is sensed separately and forms part of the signal. Thus a color CCD will contain four cells per "pixel," one reacting to red, two reacting to green (our eyes are most sensitive to green), and one reacting to blue. In some higher quality imaging systems, each color is sensed on a separate array, but a single array and filters arrangement is more economical.

CCD devices are the core of electronic cameras and video recorders but are also used in scanners (where linear arrays are typically used). They are also used for very low light applications by cooling the CCDs to low temperatures. Under these conditions their sensitivity is much higher, primarily due to reduced thermal noise. In this mode, CCDs have successfully displaced photomultipliers. An integrated CCD used in a video camera (500 lines, 625 pixels per line with four sensors [three colors] per pixel) is shown in **Figure 4.24**.

EXAMPLE 4.9 Some considerations in CCD imaging

Coupled charge devices are very common in still and video cameras, as they are inexpensive compared with other imaging devices and can be produced on a chip, making them ideal for miniaturization. The resolution, usually defined as number of pixels, can be very high while maintaining a small surface area, something that allows the use of small lenses with minimal motion for focusing and zooming. In the extreme, the imaging area can be quite large, with many millions of pixels. However, because even a simple camera contains an imaging sensor of a few million pixels, the transfer of images is not a mere formal process. Rather it is a limiting process that defines, for example, how quickly an image can be recorded.

Consider a digital camera imager with 12 megapixels in a 4:3 format. The image has 3000 columns with 4000 rows (see **Figure 4.23b** for the schematic structure). With three steps per row, the transfer of 3000 rows will take 9000 steps. Each row is first transferred into a 4000-position register that must be shifted out one cell at a time to generate the row signal. Assuming that each operation takes the same amount of time, there are $3000 \times 3 \times 4000 = 36 \times 10^6$ steps to be performed before the image is retrieved from the CCD. Assuming, arbitrarily that a step can be performed in 50 ns, it will take a minimum of 1.8 s to do so.

This means of course that a camera cannot record video in that resolution. Digital cameras that record video do so in a reduced resolution format, typically in video graphics array (VGA) or high-definition (HD) formats. For example, in VGA format the camera only records 640×480 pixels/frame. With the three steps needed, the transfer takes $640 \times 3 \times 480 = 921{,}400$ steps. At the same 50 ns step, this takes 46 ms and allows 21 frames/s, sufficient for good quality video.

Of course, many methods can and are being used to improve performance, but these simple and practical considerations give an idea of the issues involved. It should be noted as well that much higher clock speeds are not very practical, as the power consumption increases linearly with frequency, something that must also be taken into considerations in small, battery-operated cameras.

There are cameras and imaging systems that may include hundreds of megapixels. In these systems extraction of the image may take considerable time, but the quality and resolution are superior.

4.8 | THERMAL-BASED OPTICAL SENSORS

The thermal effects of radiation, that is, conversion of radiation into heat, are most pronounced at lower frequencies (longer wavelengths) and are therefore most useful in the IR portion of the spectrum. In effect, what is measured is the temperature associated with radiation. The sensors based on these principles carry different names—some traditional, some descriptive. Early sensors were known as pyroelectric sensors (from the Greek *πυρ*, “fire”). Bolometers (from the Greek *bolé*, “ray,” so bolometer can be loosely translated as a ray meter) are also thermal radiation sensors that may take various forms, but all include an absorbing element and a temperature sensor in one form or another. Some are essentially thermistors and can be used in the whole range of radiation, including microwave and millimeter wave measurements. Bolometers date back to 1878 and were originally intended to sense low light levels from space. Other names like passive infrared (PIR) and active far infrared (AFIR) not only are more descriptive, but also broader, and encompass many types of sensors.

In effect, almost any temperature sensor may be used to measure radiation as long as a mechanism can be found to transform radiation into heat. Since most methods of temperature sensing were described in **Chapter 3**, we will discuss here the specific arrangements used to sense radiation and will view the temperature-sensing elements used in conjunction with various thermal radiation sensors as given and known.

In general, thermal radiation sensors are divided into two classes: PIR and AFIR. In a passive sensor, radiation is absorbed and converted to heat. The temperature increase is measured by a sensing element to yield an indication of the radiative power. In an active sensor, the device is heated from a power source and the variations in this power due to radiation (e.g., the current needed to keep the temperature of the device constant) give an indication of the radiation.

4.8.1 Passive IR Sensors

A PIR sensor has two basic components: an absorption section that converts radiation into heat and a proper temperature sensor that converts heat into an electrical signal. Without going into the questions of heat transfer and heat capacities (these were discussed briefly in **Chapter 3**), the absorption section of the sensor must be able to both absorb as much of the incoming radiated power at the sensor's surface as possible while at the same time quickly respond to changes in radiated power density. Typically the absorber is made of a metal of good heat conductivity (gold is a common choice in high-quality sensors) that is often blackened to increase absorption. The volume of the absorber is kept small to improve response (quick heating/cooling) to changes in radiation and hence keep the response time reasonable. Typically the absorber and the sensor are encapsulated or placed in a gas-filled or evacuated hermetic chamber to avoid variations in sensing signals due to the cooling effects of air motion. The absorber is located behind a transparent (to IR radiation) window, often made of silicon, but other materials may be used (germanium, zinc selenide, etc.). The choice of the sensor materials and structure dictates to a large extent the sensitivity, spectral response, and physical construction of the device.

4.8.1.1 Thermopile PIR

In this type of device, sensing is done by a thermopile. A thermopile is made of a number of thermocouples connected in series electrically but in parallel thermally (i.e., they are exposed to identical thermal conditions). Based on the thermoelectric effect, a thermocouple generates a small potential across a junction made of two different materials. Any two materials can be used, but some material combinations produce higher potential differences (see **Section 3.3**). Thermocouples can only measure temperature differences, hence the thermopile is made of alternating junctions, as shown in **Figure 4.25**. All "cold" junctions are held at a known (measured) lower temperature, while all "hot" junctions are held at the sensing temperature. In practical construction the cold junctions are placed on a relatively large frame that has a high thermal capacity and hence the temperature will fluctuate slowly while the hot junctions are in contact with the absorber, which is small and has a low heat capacity (**Figure 4.25**). In addition, the frame may be cooled, or a reference sensor may be used on the frame so that the temperature difference can be properly monitored and related to the radiated power

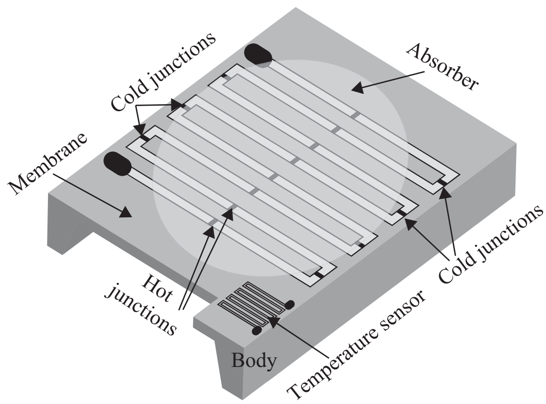


FIGURE 4.25 ■ The structure of a PIR sensor showing the thermopile used to sense temperature (under an IR absorber). A temperature sensor monitors the temperature of the cold junctions.

density at the sensor. Although any pair of materials may be used, in most PIRs crystalline or polycrystalline silicon and aluminum are used because silicon has a very high thermoelectric coefficient and is compatible with other components of the sensor, whereas aluminum has a low temperature coefficient and can be easily deposited on silicon surfaces. Other materials used (mostly in the past) are bismuth and antimony. The output of the thermocouple is the difference between the Seebeck coefficients of silicon and aluminum (see **Section 3.3**).

PIR sensors are used to sense mostly near-IR radiation but within this range they are quite common. When cooled they can be used further down into the far-IR radiation. One of the most common applications of PIR sensors is in motion detection (in which the transient temperature caused by motion is detected). However, for this purpose the pyroelectric sensors in the following section are often used because they are both simpler and less expensive than the structure described above.

EXAMPLE 4.10 Thermopile sensor for IR detection

An IR sensor is used to detect hot spots in forests to prevent fires. The sensor is made as in **Figure 4.25**, with 64 pairs of junctions using a silicon/aluminum junction (see **Tables 3.3** and **3.9**). The silicon/aluminum thermocouple has a sensitivity $S = 446$ mV/°C. The absorber is made of a thin gold foil, 10 μm thick and 2 cm^2 in area, coated black to increase absorption. Gold has a density $\rho = 19.25$ g/cm^3 and a heat capacity $C = 0.129$ J/g/°C (i.e., the absorber needs to absorb 0.129 J/g for its temperature to rise 1°C). The absorber is not ideal and its conversion efficiency is only 85% (i.e., 85% of the incoming heat is absorbed). We will denote efficiency as $e = 0.85$. The window of the sensor is $A = 2$ cm^2 and we will assume that it takes the sensor $t = 200$ ms to reach thermal steady state (i.e., for the temperature of the absorber to stabilize to a constant value at a given radiation power density). This temperature is due to the absorbed heat and heat loss. If a temperature difference between the hot and cold junction of 0.1°C can be reliably measured in the sensor, calculate the sensitivity of the sensor assuming the input is the power density of the IR radiation.

Solution: Because the input is a power density (i.e., measured in W/m^2 or in lux) and the temperature increase is due to heat (energy), the heat capacity is therefore the product of power density and time. Given a power density P_{in} [W·m^2], the power received by the sensor is the product of power density and the area of the absorber. With an 85% efficiency, the heat absorbed over 200 ms is

$$w = P_{in}tAe = P_{in} \times 0.2 \times 2 \times 10^{-4} \times 0.85 = 3.4 \times 10^{-5} P_{in} \quad [\mathrm{J}].$$

To find the temperature increase of the absorber due to this heat, we divide it by the heat capacity of the foil, denoted as C_f. This is simply the specific heat multiplied by the mass of the absorber. The latter is

$$m = 10 \times 10^{-6} \times 2 \times 10^{-4} \times 19.25 = 3.85 \times 10^{-5} \mathrm{g}.$$

The specific heat for the absorber is therefore

$$C_f = C_m = 0.129 \times 3.85 \times 10^{-5} = 4.96665 \times 10^{-6} \qquad \left[\frac{\mathrm{J}}{\mathrm{K}}\right]$$

We obtain the temperature increase in the absorber by dividing the incoming heat by C_f:

$$T = \frac{w}{C_f} = \frac{P_{in}tAe}{C_f} = \frac{3.4 \times 10^{-5}}{4.96665 \times 10^{-6}} P_{in} = 6.846 P_{in} \quad [\mathrm{K}].$$

Since the lowest measurable change in temperature is 0.1°C, we get the power density needed to raise the temperature by that amount:

$$P_{in} = \frac{0.1}{6.846} = 1.46 \times 10^{-2}\ \mathrm{W/m^2}.$$

This is 14.6 mW/m^2.

The sensitivity, by definition (assuming, of course, a linear transfer function), is the output divided by input. We have the input power density. Now we need to calculate the output of the thermocouple for that same temperature difference of 0.1°C. Since we have a thermopile with 64 pairs of junctions and a sensitivity of 446 mV/°C, the output for a temperature difference of 0.1°C is

$$V_{out} = 446 \times 64 \times 0.1 = 2854.4\ \mathrm{mV}.$$

Thus the sensitivity of the sensor is

$$S = \frac{V_{out}}{P_{in}} = \frac{2854.4}{1.46 \times 10^{-2}} = 1.955 \times 10^{5}\ \mathrm{mV/W/m^2}.$$

In practical terms this means that, for example, the sensor will produce a 1.955 V output at an input power of 10^{-5} W/m^2. This kind of sensitivity is sufficient for most low-power sensing ranging from astronomy applications to highly sensitive motion sensing.

4.8.1.2 Pyroelectric Sensors

The pyroelectric effect is an electric charge generated in response to heat flow through the body of a crystal. The charge is proportional to the change in temperature and hence the effect may be viewed as heat flow sensing rather than temperature sensing. However, in the context of this section, our interest is in the measurement of radiation and thus pyroelectric sensors are best viewed as sensing changes in radiation. For this reason they have found applications in motion sensing in which the background temperature is not important—only that due to the motion of a "warm" source is sensed. Pyroelectricity was formally named in 1824 by David Brewster, but its existence in tourmaline crystals was described in 1717 by Louis Lemery. It is interesting to note that the effect is described in the writing of Theophrastus in 314 BC as the attraction of bits of straw and ash to tourmaline when the latter was heated. The attraction is due to the charge generated by the heat. As early as the end of the nineteenth century, pyroelectric sensors were made of Rochelle salt (potassium sodium tartrate [$KHC_4H_4O_6$]). Currently there are many other materials used for this purpose, including barium titanate oxide ($BaTiO_3$), lead titanate oxide ($PbTiO_3$), as well as lead zirconium titanate (PZT) materials ($PbZrO_3$), polyvinyl fluoride (PVF), and polyvinylidene fluoride (PVDF). When a pyroelectric material is exposed to a temperature change ΔT, a charge ΔQ is generated as

$$\Delta Q = P_Q A \Delta T \quad [\mathrm{C}], \tag{4.25}$$

where A is the area of the sensor and P_Q is the pyroelectric charge coefficient defined as

$$P_Q = \frac{dP_s}{dT} \quad [\mathrm{C/m^2 \cdot K}] \tag{4.26}$$

and P_s is the spontaneous polarization [C/m^2] of the material. Spontaneous polarization is a property of the material related to its electric permittivity.

A change in potential ΔV develops across the sensor as

$$\Delta V = P_V h \Delta T \quad [\mathrm{V}], \tag{4.27}$$

where h is the thickness of the crystal and P_V is its pyroelectric voltage coefficient,

$$P_V = \frac{dE}{dT} \quad [\mathrm{V/m^2 \cdot K}], \tag{4.28}$$

and E is the electric field across the sensor. The two coefficients (voltage and charge coefficients; see **Table 4.4**) are related as follows:

$$\frac{P_Q}{P_V} = \frac{dP_s}{dE} = \varepsilon_0 \varepsilon_r \quad [\mathrm{F/m}]. \tag{4.29}$$

By definition, the sensor's capacitance is

$$C = \frac{\Delta Q}{\Delta V} = \varepsilon_0 \varepsilon_r \frac{A}{h} \quad [\mathrm{F}]. \tag{4.30}$$

Hence one can write the change in voltage across the sensor as

$$\Delta V = P_Q \frac{h}{\varepsilon_0 \varepsilon_r} \Delta T \quad [\mathrm{V}]. \tag{4.31}$$

Clearly this change in voltage is linearly proportional to the change in temperature. It should be noted again that our main interest here is not in measuring the change in temperature, but rather the change in radiation that causes this change in temperature. Also to be noted is that all sensors must operate below their Curie temperature (at the Curie temperature their polarization vanishes). **Table 4.4** shows these properties for some materials used for pyroelectric sensors.

The structure of a pyroelectric sensor is quite simple. It consists of a thin crystal of a pyroelectric material between two electrodes, as shown in **Figure 4.26a**. Some sensors use a dual element, as in **Figure 4.26b**. The second element can be used as a reference by, for example, shielding it from radiation, and it is often used to compensate for

TABLE 4.4 ■ Pyroelectric materials and some of their properties

Material	P_Q [C/m^2/°K]	P_V [V/m/°K]	ε_r	Curie temperature [°C]
TGS (single crystal)	3.5×10^{-4}	1.3×10^6	30	49
$LiTaO_3$ (single crystal)	2.0×10^{-4}	0.5×10^6	45	618
$BaTiO_3$ (ceramic)	4.0×10^{-4}	0.05×10^6	1000	120
PZT (ceramic)	4.2×10^{-4}	0.03×10^6	1600	340
PVDF (polymer)	0.4×10^{-4}	0.4×10^6	12	205
$PbTiO_3$ (polycrystalline)	2.3×10^{-4}	0.13×10^6	200	470

TGS = TriGlycine Sulfate.

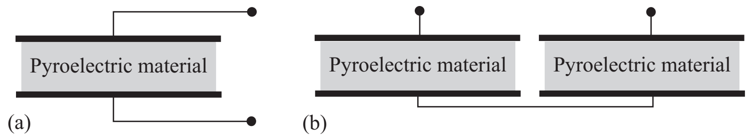


FIGURE 4.26 ■ The basic structure of a pyroelectric sensor. (a) Single element. (b) Dual element in series connected in a differential mode.

common mode effects such as vibrations or very rapid thermal changes, which can cause false effects. In the figure shown here the two elements are connected in series, but they may also be connected in parallel.

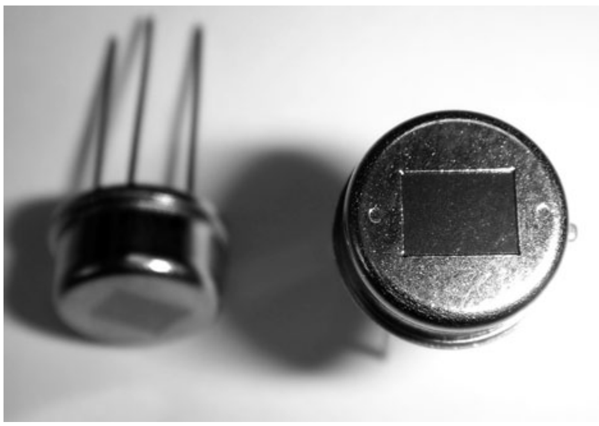
The most common materials in pyroelectric sensors are triglycine sulfate (TGS) and lithium tantalite crystals, but ceramic materials and, more recently, polymeric materials are also commonly used.

In applications of motion detection, especially of the human body (sometimes of animals), the change in temperature of IR radiation (between 4 and 20 μm) causes a change in the voltage across the sensor, which is then used to activate a switch or some other type of indicator.

An important property of all pyroelectric sensors is the decay time, during which the charge on the electrodes diffuses. This is on the order of 1–2 s because of the very high resistance of the materials, but it also depends on the external connection of the device. This response time is very important in the ability of the sensors to detect slow motion.

Figure 4.27 shows a dual IR sensor used for motion detection. This device includes a differential amplifier, operates at 3–10 V, and has a field of view of 138° horizontally (wide dimension of the window) and 125° vertically. The device has an optical bandwidth (sensitivity region) between 7 and 14 μm (in the near IR region).

FIGURE 4.27 ■ A PIR motion detector sensor. This is a dual sensor. Note the metal package and the window (3 mm × 4 mm).



EXAMPLE 4.11 Motion sensor

A motion sensor based on a PZT ceramic is used to turn on lights in a room as a person enters the room. The sensor is made of two conducting plates with a PZT chip (8 mm wide × 10 mm long × 0.1 mm thick) between them, forming a capacitor. One plate is exposed to the motion, whereas the other is connected to the body of the sensor and held at its temperature. As the person enters the room, the person's body temperature causes the exposed plate's temperature to temporarily rise by 0.01°C because of the IR radiation produced by the body. This temperature dissipates and eventually both plates will reach the same steady-state temperature. For this reason the sensor can detect motion but not presence. Calculate the charge produced on the plate and the potential difference across the sensor due to the rise in temperature.

Solution: The charge produced by the rise in temperature can be calculated from **Equation (4.25)** and the potential difference can be calculated from **Equation (4.31)**. However, we will start with **Equation (4.31)**, calculating the potential and then the charge from the relation between charge and capacitance (**Equation (4.30)**).

The potential across the plates is

$$\Delta V = P_Q \frac{h}{\varepsilon_0 \varepsilon_r} \Delta T = 4.2 \times 10^{-4} \times \frac{0.1 \times 10^{-3}}{1600 \times 8.854 \times 10^{-12}} \times 0.01 = 0.0296 \text{ V}.$$

This is a small voltage, but because the reference voltage (i.e., the output in the absence of motion) is zero, the small output voltage is easily measurable.

The charge produced by the change in temperature depends on the capacitance. The latter is

$$C = \frac{\varepsilon_0 \varepsilon_r A}{h} = \frac{1600 \times 8.854 \times 10^{-12} \times 0.008 \times 0.01}{0.1 \times 10^{-3}} = 1.1333 \times 10^{-8} \text{ F}.$$

The charge produced is

$$\Delta Q = C \Delta V = 1.1333 \times 10^{-8} \times 0.0296 = 3.355 \times 10^{-10} \text{ C}.$$

Clearly, to produce a useful output (i.e., to turn on a relay or an electronic switch) the output must be amplified. For example, if the output requires (typically) 5 V, then amplifying the voltage by about 170 will produce the required output. We will discuss these issues in **Chapter 11**, where we will see that in fact the amplification can be done using a charge amplifier. The main reason for that approach, rather than performing a classical voltage amplification, is that the impedance of pyroelectric sensors is very high, whereas the input impedance of conventional voltage amplifiers is much lower. The charge amplifier, which has a very high input impedance, is better suited for this application.

4.8.1.3 Bolometers

Bolometers are very simple radiation average power sensors useful over the whole spectrum of electromagnetic radiation, but they are most commonly used in microwave and far-IR ranges. They consist of any temperature-measuring device, but usually a small RTD or a thermistor is used. The operation is as follows: The radiation is absorbed by the device directly, causing a change in its temperature. This temperature increase is proportional to the radiated power density at the location of sensing. This change causes a change in the resistance of the sensing element that is then related to the power or power density at the location being sensed. Although there are many variations of this basic device, they all operate on essentially the same principle. However, since the temperature increase due to radiation is the measured quantity, it is important that the background temperature (i.e., air) be taken into account. This can be done by separate measurements or by a second bolometer that is shielded from the radiation (usually in a metal enclosure in the case of microwaves). The sensitivity of a bolometer to radiation can be written as follows:

$$\beta = \frac{\alpha \varepsilon_s}{2} \sqrt{\frac{Z_T R_0 \Delta T}{(1 + \alpha_0 \Delta T)[1 + (\omega \tau)^2]}}, \tag{4.32}$$

where $\alpha = (dR/dT)/R$ is the temperature coefficient of resistance (TCR) of the bolometer, ε_s is its surface emissivity, Z_T is the thermal resistance of the bolometer, R_0 is its resistance at the background temperature, ω is the frequency, τ is the thermal time constant, and ΔT is the increase in temperature. Clearly then the ideal bolometer should have a large resistance at background temperatures and high thermal resistance. On the other hand, they must be small physically, which favors low thermal resistances.

In terms of construction, bolometers are fabricated as very small thermistors or RTDs, usually as individual components or as integrated devices. In all cases it is important to insulate the sensing element from the structure supporting it so that its thermal impedance is high. This can be done by simple suspension of the sensor by its wires or, as is sometimes done, by suspension of the sensor over a silicon groove.

Bolometers are some of the oldest devices used for this purpose and have been used for many applications in the microwave region, including mapping of antenna radiation patterns, detection of IR radiation, testing of microwave devices, and much more.

4.9 | ACTIVE FAR INFRARED (AFIR) SENSORS

In its simplest form an active infrared (AFIR) sensor can be thought of as a power source that heats the sensing element to a temperature above ambient and keeps its temperature constant. When used to sense radiation, additional heat is provided to the sensors through this radiation. The power necessary to keep the temperature constant is now reduced and the difference in power is a measure of the radiation power. Of course, in practice the process is more complicated. Assuming that the temperature of the sensing element is constant, the AFIR sensor can be viewed as being time independent. Under these conditions the power supplied to the sensor through an electric circuit that heats it up to a constant temperature T_s is

$$P = P_L + \Phi \quad [\mathrm{W}], \tag{4.33}$$

where $P = V^2/R$ is the heat supplied by a resistive heater (V is the voltage across the heating element and R its resistance), P_L is the power loss through conduction through the body of the sensor, and Φ is the radiation power being sensed.

The power loss is

$$P_L = \alpha_s (T_s - T_a) \quad [\mathrm{W}], \tag{4.34}$$

where α_s is a loss coefficient or thermal conductivity (which depends on materials and construction), T_s is the sensor's temperature, and T_a is the ambient temperature. Given the power supplied as $P = V^2/R$, a sensor with surface area A, and a total emissivity ε, the sensed temperature, T_m, is

$$T_m = \sqrt[4]{T_s^4 - \frac{1}{A\sigma\varepsilon}\left[\frac{V^2}{R} - \alpha_s(T_s - T_a)\right]}, \tag{4.35}$$

where σ is the electric conductivity of the sensor medium.

This relation gives the temperature as a function of voltage across the heating element. By measuring this voltage, a reading of the radiation power is obtained.

Although AFIR devices are much more complex than simple PIRs, including bolometers, they have the advantage of a much higher sensitivity and an independence from thermal noise that other IR sensors do not possess. Hence AFIR devices are used for low-contrast radiation measurements where PIRs are not suitable.

4.10 | OPTICAL ACTUATORS

When discussing optical devices it is not immediately clear what an optical actuator might be because of our tendency to think of actuators as devices that perform some kind of motion. However, based on our definition of actuators, in fact there are many optical actuators and they are quite common. The use of a laser beam to perform eye surgery, to machine a material, or to record data in a magneto-optical hard drive are some examples. Others are transmission of data on an optical fiber, transmission of a command using an IR remote control device, use of an LED or laser to illuminate a CD to read the data, scanning of an IPC code in a supermarket, or even turning on a light in a room. In **Chapter 10** we will discuss additional examples such as optical switching.

Optical actuation can be low or high power. In an optical link such as in optical fiber communication or in an optical isolator (to be discussed in **Chapter 12**), the transmitting element (the actuator) is a low-power LED in the IR or visible range (**Figure 4.28**). The power produced by the LED may only be a few milliwatts. On the other hand, industrial lasers, such as carbon dioxide (CO_2) lasers (in which CO_2 gas is excited and produces a beam in the IR region around 100 μm), can produce hundreds of kilowatts of useful power for a variety of industrial processing purposes, including machining, surface treatment, and welding. In between are lasers, mostly CO_2 lasers of moderate power (a few watts to a few hundred watts) used for medical applications, including surgery, skin ablation, and suturing. Other applications are for range finding, particularly in the military, and for speed detection and measurement. Laser actuation is also used for the production of electronic components, for trimming of devices, and even for recording of data on CD-ROMs, where they actually scribe the surface with the pattern representing the data.

A good example of an optical actuator of the type described here is the magneto-optical recording of data, a common method used for high-density data recording. The principle is shown schematically in **Figure 4.29** and is based on two principles. First, a laser beam, focused to a small point, can heat the surface of the disk to a high temperature in a few nanoseconds. Second, when a ferromagnetic material such as iron or its oxides is heated above a certain temperature (about 650°C), the material loses

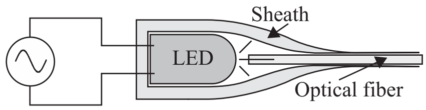


FIGURE 4.28 ■ An optical link. The fluctuations in the light intensity of the LED represent the data transmitted along the optical fiber.

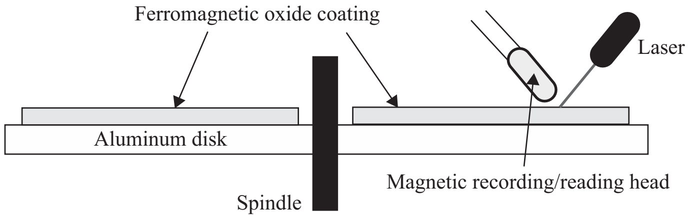


FIGURE 4.29 ■ Magneto-optical recording. The laser beam heats the recording medium above the Curie temperature and the magnetic recording head applies the magnetic field needed to record the data.

its magnetic properties. This temperature is called the **Curie temperature** and is characteristic of the particular material used as the recording medium (mostly Fe_2O_3). When cooled, the material becomes magnetized with the field supplied by the recording head. To record data, the laser is turned on to heat the point above the Curie temperature and the datum that needs to be recorded at that point is supplied in the form of a low-intensity magnetic field by a magnetic recording head. The beam is then switched off and the spot cools below the Curie temperature in the presence of the magnetic field, retaining the data permanently. Erasure of data is done by heating the spot and cooling it off without a magnetic field. The data are read using the magnetic recording head alone. The advantage of this method is that the data density is much higher than purely magnetic recording, which requires larger magnetic fields that in turn extend over larger surfaces and hence is only practical at lower data densities.

4.11 | PROBLEMS

Optical units

4.1 **Optical quantities.** An isotropic light source (a source that radiates uniformly in all directions in space) produces a luminance of 0.1 cd/m^2 at a distance of 10 m. What is the power of the source?

4.2 **Optical sensitivity.** Many optical instruments such as video cameras are rated in terms of sensitivity in lux, especially to indicate low-light sensitivity. One may encounter specifications such as "sensitivity – 0.01 lux." What is the sensitivity in terms of power density?

4.3 **Electron volts and joules.** The Planck constant is given as 6.6261×10^{-34} J/s or as 4.1357×10^{-15} eV/s. Show that the two quantities are identical but expressed in different units.

The photoelectric effect

4.4 **Photon energy and electron kinetic energy.** A photoelectric device intended for UV sensing is made of a platinum cathode. Calculate the range of kinetic energies of the electrons emitted by UV light between 400 nm and 1 pm, assuming that a photon emits a single electron.

4.5 **Work function and the photoelectric effect.** The work function for copper is 4.46 eV.

a. Calculate the lowest photon frequency that can emit an electron from copper.

b. What is the wavelength of the photons and in what range of optical radiation does that occur?

4.6 **Electron density in photoelectric sensor.** A photoelectric sensor has a cathode in the form of a disk of radius $a = 2$ cm coated with an alkali compound that has a work function $e_0 = 1.2$ eV and a quantum efficiency of 15%. The sensor is exposed to sunlight. Calculate the average number of electrons emitted per second assuming a power density of 1200 W/m^2 and uniform distribution of power over the spectrum between red (700 nm) and violet (400 nm).

4.7 **Work function, kinetic energy, and current in a photoelectric sensor.** A photoelectric sensor with unknown cathode material is subjected to experimental evaluation. The current in the sensor (see **Figure 4.19**) is measured while the wavelength of the radiation is recorded. Emission is observed starting with red light at 820 nm. The power density of the incoming radiation is kept constant at 5 mW/cm^2.

a. Calculate the work function of the cathode.

b. If the photoelectric sensor is now illuminated with the same power density, but with a blue light at a wavelength of 480 nm, what is the kinetic energy of the electrons released?

c. Assume each photon releases one electron. Calculate the current in the sensor if the cathode has an area of 2 cm^2.

The photoconducting effect and photoconducting sensors

4.8 **Bandgap energy and spectral response.** A semiconductor optical sensor is required to respond down to 1400 nm for use as a near-IR sensor. What is the range of bandgap energies of the semiconductor that can be used for this purpose?

4.9 **Germanium silicon and gallium arsenide photoconducting sensors.** The intrinsic concentrations and mobilities for germanium, silicon, and gallium arsenide are as follows:

	Germanium (Ge)	Silicon (Si)	Gallium Arsenide (GaAs)
Intrinsic concentration [per cm^3]	2.8×10^{13}	1.0×10^{10}	2.0×10^{6}
Mobility of electrons, μ_e [cm^2/V/s]	3900	1400	8800
Mobility of holes, μ_p [cm^2/V/s]	1900	450	400

Compare the dark resistance of an identical sensor for the three materials made as a rectangular bar of length 2 mm, width 0.2 mm, and thickness 0.1 mm to be used as photoconductors. The resistance calculated here is the nominal resistance of the sensor.

4.10 Gallium arsenide photoconductive sensor. A gallium arsenide (GaAs) photoconductive sensor is made as a small, rectangular chip 2 mm long, 1 mm wide, and 0.1 mm thick (see **Figure 4.4a** for the construction). A red light of intensity 10 mW/cm^2 and wavelength 680 nm is incident perpendicularly on the top surface. The semiconductor is *n*-type with an electron concentration of 1.1×10^{19} electrons/m^3. The mobilities of electrons and holes in GaAs are 8400 cm^2/V·s and 400 cm^2/V·s, respectively. Assuming all incident power on the top surface of the device is absorbed and a quantum efficiency of 0.38, calculate

a. The "dark" resistance of the sensor.

b. The resistance of the sensor when light shines on it. The recombination time of electrons is approximately 10 μs.

4.11 Improved gallium arsenide photoconductive sensor. In an attempt to improve the sensor described in **Problem 4.10**, the meander shape in **Figure 4.4b** is adopted, keeping the total exposed area the same (2 mm^2) but reducing the length between the electrodes to 0.5 mm.

a. Calculate the "dark" resistance of the sensor.

b. Calculate the resistance of the sensor when light shines on it.

c. Compare the results with those for the rectangular sensor and comment on the sensitivity of the two devices.

4.12 Intrinsic silicon optical sensor. Suppose a photoconductive sensor is made of intrinsic silicon with structure and dimensions as shown in **Figure 4.30**. The intrinsic concentration is 1.5×10^{10} carriers/cm^3 and mobilities of electrons and holes are 1350 cm^2/V·s and 450 cm^2/V·s, respectively. The carrier lifetime for electrons and holes depends on the concentration and changes with illumination, but for simplicity we will assume these are constant at 10 μs. Assume also that 50% of the incident power is absorbed by the silicon and 45% quantum efficiency.

a. Find the sensitivity of the device to input power density at a given wavelength in general terms.

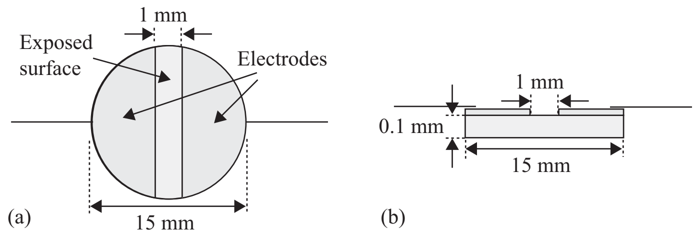


FIGURE 4.30 ■ A photoconducting sensor.

b. What is the sensitivity at 1 mW/cm^2 at a wavelength of 480 nm?

c. What is the cutoff wavelength, that is, the wavelength beyond which the sensor cannot be used?

Photodiodes

4.13 Photodiode in photoconductive mode. A photodiode is connected in reverse mode with a small reverse voltage to ensure low reverse current. The leakage current is 40 nA and the sensor operates at 20°C. The junction has an area of 500 μm^2 and operates from a 3 V source, as shown in **Figure 4.31**. The resistor is 240 Ω. Calculate the voltage across R in the dark and when illuminated with a red laser beam (800 nm) with power density of 5 mW/cm^2. Assume a quantum efficiency of 50%.

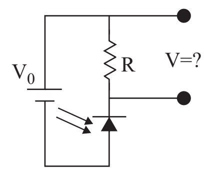


FIGURE 4.31 ▪ A photodiode in photoconductive mode.

4.14 Photodiode in forward bias. A photodiode is forward biased as shown in **Figure 4.32** so that the voltage across the diode is 0.2 V. The dark current through the diode is 10 nA.

a. What must be the voltage V_0 to obtain this bias at 20°C in the dark?

b. When illuminated with a certain power density, and with the voltage V_0 found in (a) applied, the bias changes to 0.18 V. Calculate the total power absorbed by the diode at 800 nm.

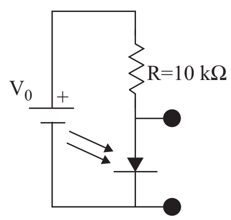


FIGURE 4.32 ▪ Forward-biased photodiode.

4.15 Dusk/dawn light switch. Many lighting systems, including street lightning, turn on and off automatically based on light intensity. To do so, it is proposed to use a photodiode in the configuration shown in **Figure 4.33**. The diode has negligible dark current, an exposed area of 1 mm^2, and a quantum efficiency of 35%. The electronic switch is designed so that to turn on the lights the voltage across R must be 8 V or less and to turn off the lights the voltage must be 12 V or greater. On a normal sunny day, the power density available at ground level is 1200 W/m^2.

a. Calculate the resistance R so that the lights turn on when available light intensity (in the evening) is 10% (or less) of the normal daylight. Assume radiation at an average wavelength of 550 nm.

b. At what light intensity will the lights turn off in the morning?

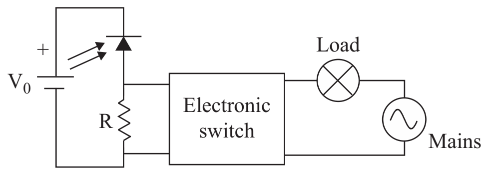


FIGURE 4.33 ■ A light-activated switch.

c. Repeat (a) and (b) if it is known that in the evening the average light wavelength tends to be more red, with an average wavelength of 580 nm, and in the morning it tends to be more blue, with an average wavelength of 520 nm.

Photovoltaic diodes

4.16 **Solar cells as actuators: power generation.** The use of solar cells for power generation is common in small-scale installations and for use in stand-alone equipment such as remote sensors and monitoring stations. To get some idea of what is involved, consider a solar cell panel with an overall power conversion efficiency of 30% (i.e., 30% of the power available at the surface of the solar cell is converted into electrical power). The panel is 80 cm $\times$ 100 cm in area and the maximum solar power density at the location is 1200 W/m^2. The panel is divided into 40 equal-size cells and the cells are connected in series. Assume that the cells are equally responsive over the entire visible spectrum (400–700 nm), have a quantum efficiency of 50%, and the internal resistance of the 40 cells in series is 10 Ω. The leakage current of each cell is 50 nA. Calculate the maximum power the solar cell can deliver and indicate what the conditions must be for that to happen.

4.17 **Overall efficiency of solar cells.** A solar panel supplies a current of 0.8 A into a 10 Ω load when exposed to the sun under optimal conditions (i.e., the sun's radiation is vertical, power density is maximal) at a location where the sun's radiation intensity is 1400 W/m^2. The panel is made of cells, each 10 cm $\times$ 10 cm, and the cells are connected in series. Use the average wavelength in the visible range of 550 nm as the wavelength of radiation.

a. Calculate the overall conversion efficiency of the solar cells under the stated conditions.

b. What is the maximum power the cell can deliver into a 10 Ω load if its overall efficiency can be increased to 30%? Assume the internal resistance is 10 Ω.

Phototransistors

4.18 **Phototransistor as a detector.** Consider again **Example 4.4**, but now the photodiode is replaced with a phototransistor biased as shown in **Figure 4.34**. The

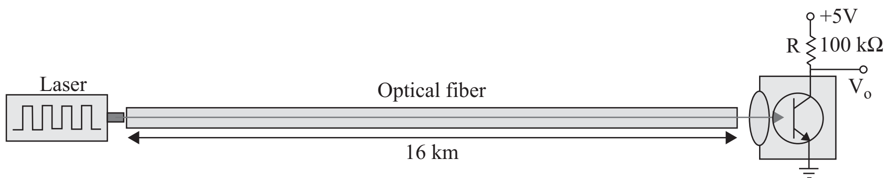


FIGURE 4.34 ■ Phototransistor as a detector in an optical link.

phototransistor has a gain of 50. For a given input pulse train, show the output in relation to the input and calculate the voltage levels expected. Assume that all power that reaches the phototransistor is absorbed in the base–emitter junction.

4.19 Phototransistor and saturation current. The phototransistor in **Figure 4.35** operates in its linear range. At an input power density of 1 mW/cm^2, the voltage measured between the collector and emitter is $V_{ce} = 10.5$ V. Calculate the span of the phototransistor as a light intensity sensor. That is, calculate the minimum light intensity (for which the current through the transistor is zero) and maximum power density (at which the current through the transistor saturates). The saturation V_{ce} voltage is 0.1 V. Neglect the effect of the dark current.

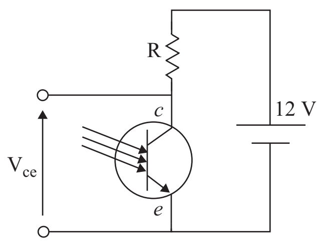


FIGURE 4.35 ■ A phototransistor and its operation.

4.20 Temperature effects on phototransistors. Consider again the configuration in **Figure 4.35**, where $R = 1$ kΩ and the amplification (gain) of the transistor is 100. The collector–emitter voltage at a light intensity of 1 mW/cm^2 is equal to 8 V at 20°C. If the light is removed, the collector–emitter voltage climbs to 11.8 V. Calculate the collector–emitter voltage if the temperature rises to 50°C. Assume that V_{BE} does not change with temperature or with light intensity and the dark current is 10 nA. Discuss the result.

Note: V_{be} decreases with temperature at a rate of 1.0–2.0 mV/°C (see **Section 3.4**), but we will neglect this change here.

Photoelectric sensors and photomultipliers

4.21 Current and electron velocity in a photomultiplier. To get some idea of the processes occurring in a photomultiplier, consider the following simplified configuration. A circular cathode and a circular anode each of radius $a = 20$ mm are separated $d = 40$ mm apart and connected to a potential difference $V = 100$ V as shown in **Figure 4.36**. The cathode is made of a potassium-based compound with a work function $e_0 = 1.6$ eV and a quantum efficiency $\eta = 18\%$.

a. If blue light at a wavelength of 475 nm and intensity of 100 mW/cm^2 impinges on the cathode, calculate the current expected in the device.

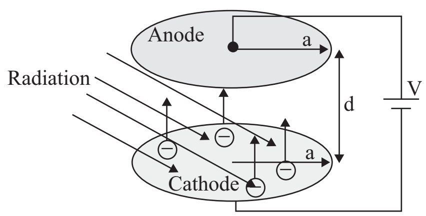


FIGURE 4.36 ■ A basic photomultiplier.

b. Calculate the velocity of the electrons when they reach the anode, given the mass of the electron $m_e = 9.1094 \times 10^{-31}$ kg.

4.22 **Limit sensitivity of a photomultiplier in the visible range.** Given a photomultiplier, its highest sensitivity occurs at the shorter wavelengths. Assume that the wavelength is 400 nm (violet light). If the cathode has a work function $e_0 =$ 1.2 eV and quantum efficiency of 20%, calculate the lowest illuminance the photomultiplier can discern if detection requires emission of at least 10 electrons/mm^2 of the cathode area per unit time.

CCD sensors and detectors

4.23 **Digital video camera image transfer from CCD.** A digital color video camera requires a picture format of 680 pixels $\times$ 620 pixels for display on a TV screen. Assuming that smooth video requires 25 frames/s (PAL or SECAM formats), calculate the minimum clock frequency for the stepping process necessary to retrieve the image from the CCD. Neglect the time needed to create the image.

4.24 **CCD sensor for HD video image transfer.** The CCD sensor in an HD video recorder is arranged as 1080 lines with 1920 pixels/line and produces 60 frames/s. Calculate the minimum clock frequency needed to transfer the image assuming the basic process in **Figure 4.23b** is used (there are other methods and modifications that can be used).

Thermopile PIR sensors

4.25 **Thermopile PIR sensor.** An IR sensor designed to operate at high temperatures is made as in **Figure 4.25** with 32 pairs of carbon/silicon carbide junctions (see **Chapter 3, Table 3.4**). The carbon/silicon carbide thermocouple has a sensitivity $s = 170$ mV/°C. The absorber is made of thin tungsten foil, 10 μm thick and 2 cm^2 in area, coated black to increase absorption. Tungsten has a density of 19.25 g/cm^3 (same as gold) and a heat capacity $C = 24.27$ J/mol/K. The absorber's conversion efficiency is 80%. The window of the sensor is $A = 5$ cm^2 and it takes the sensor $t = 300$ ms to reach thermal steady state, that is, for the temperature of the absorber to stabilize to a constant value at a given radiation power density. If a temperature difference between the hot and cold junction of 0.5°C can be reliably measured in the sensor, calculate the sensitivity of the sensor.

4.26 **Thermopile IR sensor.** An IR sensor is required to develop an output of 5 V for an IR radiation of 20 mW/cm^2. For increased sensitivity, aluminum/silicon thermocouples are used because they have an output of 446 mV/°C. An aluminum absorber of area 4 cm^2 and thickness 20 μm is used with an absorption efficiency of 80%. Aluminum has a density of 2.712 g/cm^3 and a heat capacity of 0.897 J/g/°C. The required output must be obtained within 100 ms.

a. What is the increase in the temperature of the absorber?

b. Calculate the number of thermocouples needed to obtain the required output.

c. If the output of the device is measured using a digital voltmeter with a resolution of 0.1 V, what is the effective resolution of the sensor?

Pyroelectric sensors

4.27 **Pyroelectric motion sensor.** A PZT motion sensor is required to detect the motion of a body within its range. The sensor is built as a dual element (see **Figure 4.26**) to reduce the influence of temperature changes that are not due to the motion of a body.

a. Calculate the sensitivity if each element is 0.1 mm thick. One element is exposed to the IR source, the other is shielded from it.

b. Show how temperature compensation for common-source heat (heat sources that affect both element equally) is accomplished.

4.28 **Time constant in motion sensors.** A pyroelectric sensor is made of a small barium titanate oxide ($BaTiO_3$) chip, 10 mm × 10 mm in area and 0.2 mm thick, sandwiched between two metal electrodes. In addition to the properties given in **Table 4.4**, barium titanate oxide has a conductivity of 2.5×10^{-9} S/m.

a. Calculate the time constant of charge decay after heat has been removed.

b. The sensor is used as a motion sensor to automatically trigger a camera for wildlife photography at night. A cat runs by the sensor, generating a temperature differential of 0.1°C, triggering the camera. If at least half the charge generated across the plates must discharge before the sensor can retrigger the camera, how much time is needed before the next event can be detected?

c. To discharge the sensor faster, one can connect a resistor across the sensor. If the sensor must be ready to trigger within 250 ms, what must be the resistance across the sensor? What is the side effect of connecting a smaller resistor across the sensor other than a quicker retrigger time?

Optical actuators

4.29 **Coupling of power in an optical link.** Optical links are very common in data communication (see **Figure 4.10** and **Problem 4.20**). However, linking optical power to optical fibers can be a very low efficiency affair if not done properly. Consider two ways of coupling the power from an LED to an optical fiber, shown in **Figure 4.37**. In **Figure 4.37a**, the optical fiber is simply held in front of the LED, whereas in **Figure 4.37b** an intervening light guide is used.

a. If the LED radiates 20 mW uniformly over a 5° cone and the optical fiber has a diameter of 130 μm, calculate the power coupled to the optical fiber using the method in **Figure 4.37a**. Neglect any reflections that may occur at the

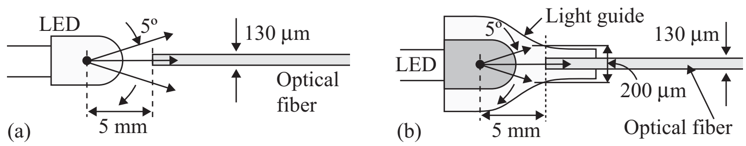


FIGURE 4.37 ■ Coupling light to optical fibers. (a) Direct coupling. (b) Use of a light guide to increase coupled power.

interface between the optical fiber and air. The distance between the LED source and surface of the fiber is 5 mm.

b. How much power is coupled in **Figure 4.37b**? Assume that all power follows the light guide with uniform power density across the light guide cross section and none can escape through its outer surface. The light guide has circular cross section as does the fiber.

4.30 **Magneto-optical recording.** In a hard disk storage device, writing data is done by magneto-optical means. Writing is done by heating the location where the data are written to the Curie temperature of the storage medium and applying the magnetic field representing the data to that spot while it cools below the Curie temperature (see **Figure 4.29**). The data are written on iron oxide (Fe_2O_3) coated on a conducting disk. Assume that 80% of the writing time is needed to heat the spot. The laser beam is 1 μm in diameter and supplies a power of 50 mW. The storage medium is 100 nm thick and has a heat capacity of 23.5 J/mol/K and a density of 5.242 g/cm^3. The curie temperature is 725°C.

a. Calculate the maximum writing data rate of the drive at an ambient temperature of 30°C.

b. Discuss possible ways the data rate can be increased.

c. Discuss effects that will reduce the data rate in practical applications.